

Schaefer L.  
09/1966/608  
Seg. ID 2 w/ Integrator

Run on: March 4, 2002, 20:24:12 ; Search time 116.27 Seconds  
 (without alignments)  
 238.804 Million cell updates/sec

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

OM protein - protein search, using sw model

Title: US-09-966-608-2

Sequence: 1 DLASAVGVIQSGSIFHHFKSK..... SLSAEGQAHVIALRVDVYEQI 100

Scoring table: BLOSUM62

Perfect score: 493

Gapext 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

- /cggn2\_6/podata/2/paa/us084\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/PCTRS\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us07\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us080\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us081\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us082\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us083\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us085\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us086\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us087\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us088\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us089\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us090\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us091\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us092\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us093\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us094\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us095\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us096\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us097\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us098\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us099\_COMBO.pep:\*
- /cggn2\_6/podata/2/paa/us60\_COMBO.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	493	100.0	1	PCP-US00-11473-2
2	493	100.0	1	PCT-US01-16032-3
3	493	100.0	1	PCT-US01-16032A-3
4	493	100.0	100	US-09-303-120B-2
5	493	100.0	100	US-09-575-061-3
6	493	100.0	100	US-09-820-516-2
7	429	87.0	245	US-09-252-931A-66292
8	267	54.2	206	US-09-328-5572
9	110.5	22.4	190	PCR-US00-11473-6

RESULT PCT-US00-11473-2

; Sequence 2, Application PC/TUS0011473

; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: IBD-Associated Microbial Antagonist

; TITLE OF INVENTION: Same

; FILE REFERENCE: FP-PM 4142

; CURRENT APPLICATION NUMBER: PCT/TUS00/11473

; CURRENT FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 09/303,120

; PRIOR FILING DATE: 1999-04-30

; NUMBER OF SEQ ID NOS.: 12

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 2

; LENGTH: 100

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism

; OTHER INFORMATION: Description from the human gut

Query Match Score: 493; DB 1;

Best Local Similarity: 100.0%; Pred. No. 4.6e-51

Matches: 100; Conservative: 0; Mismatches: 0

Oy Db

1 DLASAVGVIQSGSIFHHFKSKDELRAVMETIHNTAMRAS

1 DLASAVGVIQSGSIFHHFKSKDELRAVMETIHNTAMRAS

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

RESULT 2  
PCT-US01-16032-3  
; Sequence 3, Application PC/TUS0116032  
; GENERAL INFORMATION:  
; APPLICANT: Cedars-Sinai Medical center  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of  
FILE REFERENCE: P-PM 4713  
CURRENT APPLICATION NUMBER: PCT/US01/16032  
PRIORITY FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Microbial organism from the human gut  
PCT-US01-16032-3

Query Match 100.0%; Score 493; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.6e-51; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60  
Db 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60

QY 61 LOSIMGGSGEAMAVLVYEWRSLSAEGOAHVIALRDVYEQI 100  
Db 61 LOSIMGGSGEAMAVLVYEWRSLSAEGOAHVIALRDVYEQI 100

RESULT 3  
PCT-US01-16032A-3  
; Sequence 3, Application PC/TUS0116032A  
; GENERAL INFORMATION:  
; APPLICANT: Targan, Stephan R.  
; APPLICANT: Braun, Jonathan  
; APPLICANT: Sutton, Christopher L.  
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of  
FILE REFERENCE: FP-PW 4713  
CURRENT APPLICATION NUMBER: PCT/US01/16032A

PRIORITY FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Microbial  
US-09-303-120B-2

Query Match 100.0%; Score 493; DB 17; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.6e-51; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60  
Db 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60

QY 61 LOSIMGGSGEAMAVLVYEWRSLSAEGOAHVIALRDVYEQI 100  
Db 61 LOSIMGGSGEAMAVLVYEWRSLSAEGOAHVIALRDVYEQI 100

RESULT 5  
US-09-575-061-3  
; Sequence 3, Application US/09575061  
; GENERAL INFORMATION:  
; APPLICANT: Targan, Stephan R.  
; APPLICANT: Braun, Jonathan  
; APPLICANT: Sutton, Christopher L.  
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of  
; TITLE OF INVENTION: Crohn's Disease Using The OmpC Antigen  
FILE REFERENCE: P-PM 4097  
CURRENT APPLICATION NUMBER: US/09/575,061  
CURRENT FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Microbial organism from the human gut  
US-09-575-061-3

Query Match 100.0%; Score 493; DB 19; Length 100;  
Best Local Similarity 100.0%; Pred. No. 4.6e-51; Mismatches 0; Indels 0; Gaps 0;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60  
Db 1 DLASAVGIOQSIFHHRKSKDEILRAVMEETHNTAMMRASLEASTVRVLALIRCE 60

RESULT 6  
 US-09-820-576-2  
 ; Sequence 2, Application US/09820576  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Braun, Jonathan  
 ; APPLICANT: Sutton, Christopher L.  
 ; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid  
 ; FILE REFERENCE: P-PM 4646  
 ; CURRENT APPLICATION NUMBER: US/09/820,576  
 ; CURRENT FILING DATE: 2001-03-28  
 ; PRIOR APPLICATION NUMBER: US 09/303,120  
 ; PRIORITY FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SEQ ID NO 2 LENGTH: 100  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: Microbial organism from the human gut

Query Match 100 %; Score 493; DB 22; Length 100;  
 Best Local Similarity 100 %; Pred No. 4.6e-51; Mismatches 0; Indels 0; Gaps 0;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEASTVRVVALIRCE 60  
 Db 1 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEASTVRVVALIRCE 60

RESULT 7  
 US-09-252-991A-26292  
 ; Sequence 26292, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 26292 LENGTH: 245  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa

Query Match 87 %; Score 429; DB 16; Length 245;  
 Best Local Similarity 84.0%; Pred. No. 8.2e-43; Mismatches 6; Indels 0; Gaps 0;  
 Matches 84; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEASTVRVVALIRCE 60  
 Db 87 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEASTVRVVALIRCE 145  
 Qy 61 LOSIMCGSGEAMAVLVYEWRSLSAEGQAHVLALRDYEQI 100

RESULT 8  
 US-09-328-352-5572  
 ; Sequence 5572, Application US/09328352  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 5572 LENGTH: 206  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii

Query Match 54.2%; Score 267; DB 17; Length 206;  
 Best Local Similarity 54.0%; Pred No. 2.2e-23; Mismatches 21; Indels 0; Gaps 0;  
 Matches 54; Conservative 21; Mismatches 25; Indels 0; Gaps 0;

Qy 1 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEASTVRVVALIRCE 60  
 Db 50 ELAQFGTQGSLSFHFKSKDILAHMECTIYLNARLDAVQTSTDPEQQLRALKSE 109  
 Qy 61 LOSIMCGSGEAMAVLVYEWRSLSAEGQAHVLALRDYEQI 100  
 Db 110 LISITGDGTGAMAVLVYEWRSLSAEGQAHVLALRDYEQI 149

RESULT 9  
 PCT-US00-11473-6  
 ; Sequence 6, Application PCT/US0011473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using  
 ; TITLE OF INVENTION: Same  
 ; FILE REFERENCE: FP-PM 4142  
 ; CURRENT APPLICATION NUMBER: PCT/US00/11473  
 ; PRIOR APPLICATION NUMBER: US 09/303,120  
 ; PRIOR FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6 LENGTH: 190  
 ; TYPE: PRT  
 ; ORGANISM: Clostridium pasteurianum  
 PCT-US00-11473-6

Query Match 22.4%; Score 110.5; DB 1; Length 190;  
 Best Local Similarity 27.0%; Pred No. 0.00014; Mismatches 35; Indels 13; Gaps 3;  
 Matches 27; Conservative 25; Mismatches 35; Indels 13; Gaps 3;

Qy 1 DLASAVGIOGSISIFHFHKSDIELURAVMETIHTNTAMMRASLEAS---TVRVALIRCE 56  
 Db 28 ETASNAGVAKGTYLFHKSKEEFKVIVTEGVN---LMKNEDEATDKETALEKLKAV 83  
 Qy 57 IRCELOSIMCGSGEAMAVLVYEWRSLSAEGQAHVLALRDYEQI 96  
 Db 84 CRVOLNLYKRNDRFFKVIAQLWGR---ELROLELRDI 118

RESULT 10  
 US-09-303-120B-6  
 ; Sequence 6, Application US/09303120B  
 ; GENERAL INFORMATION:

APPLICANT: Braun, Jonathan  
 APPLICANT: Sutton, Christopher  
 TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using  
 FILE REFERENCE: P-PM 3478  
 CURRENT FILING DATE: 1999-04-30  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 190  
 TYPE: PRT  
 ORGANISM: Clostridium pasteurianum  
 US-09-303-120B-6

Query Match 22.4%; Score 110.5; DB 17; Length 190;  
 Best local Similarity 27.0%; Pred. No. 0.00014;  
 Matches 27; Conservative 25; Mismatches 35; Indels 13; Gaps 3;

Qy 1 DLASAVGIOGSISFHFKSKDEILRAVMEETHYNTAMMRASLEAS---TVRVLALIR 56  
 Db 28 EIASNAGYAKGTLYHFSKKEEFLKYIIEGVN---LMKNEIDEATDKEKTALEKLKAV 83

Qy 57 IRCLQLSIMGGSGEAMAVLYEWSLSAEGQAIWALRDV 96  
 Db 84 CRVOLNLTYKNRDFEKVITASQLWCK----ELRQELRDI 118

RESULT 11  
 US-09-820-576-6  
 Sequence 6, Application US/09820576  
 GENERAL INFORMATION:  
 APPLICANT: Braun, Jonathan  
 APPLICANT: Sutton, Christopher L.  
 TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid  
 TITLE OF INVENTION: Molecules  
 FILE REFERENCE: P-PM 4645  
 CURRENT APPLICATION NUMBER: US/09/820, 576  
 CURRENT FILING DATE: 2001-03-28  
 PRIOR APPLICATION NUMBER: US 09/303, 120  
 PRIOR FILING DATE: 1999-04-30  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 6  
 LENGTH: 190  
 TYPE: PRT  
 ORGANISM: Clostridium pasteurianum  
 US-09-820-576-6

Query Match 22.4%; Score 110.5; \*DB 22; Length 190;  
 Best Local Similarity 27.0%; Pred. No. 0.00014;  
 Matches 27; Conservative 25; Mismatches 35; Indels 13; Gaps 3;

Qy 1 DLASAVGIOGSISFHFKSKDEILRAVMEETHYNTAMMRASLEAS---TVRVLALIR 56  
 Db 28 EIASNAGYAKGTLYHFSKKEEFLKYIIEGVN---LMKNEIDEATDKEKTALEKLKAV 83

Qy 57 IRCLQLSIMGGSGEAMAVLYEWSLSAEGQAIWALRDV 96  
 Db 84 CRVOLNLTYKNRDFEKVITASQLWCK----ELRQELRDI 118

RESULT 13  
 PCT-US00-11473-7  
 Sequence 7, Application PC/TUS0011473  
 GENERAL INFORMATION:  
 APPLICANT: The Regents of the University of California  
 TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using  
 TITLE OF INVENTION: Same  
 FILE REFERENCE:  
 CURRENT APPLICATION NUMBER: PCT/US00/11473  
 CURRENT FILING DATE: 2000-04-28  
 PRIOR APPLICATION NUMBER: US 09/303, 120  
 PRIOR FILING DATE: 1999-04-30  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 7  
 LENGTH: 200  
 TYPE: PRT  
 ORGANISM: Mycobacterium tuberculosis  
 PCT-US00-11473-7

Query Match 19.2%; Score 94.5; DB 1; Length 200;  
 Best local Similarity 26.9%; Pred. No. 0.013;  
 Matches 21; Conservative 20; Mismatches 34; Indels 3; Gaps 1;

Qy 1 DLASAVGIOGSISFHFKSKDEILRAVMEETHYNTAMMRASLEAS---TVRVLALIRCE 60  
 Db 35 DIADGAGILSGLSHHFNSKKEEAVDELRLRGFLWFLARYRVDVSTANPLERQGLEMAS 94

Qy 61 LQSIMGGSGEAMAVLYE 78  
 Db 95 FEAEHHHAQQ--WVIYQ 109

RESULT 14  
 US-09-303-120B-7  
 Sequence 7, Application US/09303120B  
 GENERAL INFORMATION:  
 APPLICANT: Braun, Jonathan  
 APPLICANT: Sutton, Christopher  
 TITLE OF INVENTION: IBD-Associated Microbial Antigens and Methods of Using  
 FILE REFERENCE: P-PM 3478  
 CURRENT APPLICATION NUMBER: US/09/303, 120B  
 CURRENT FILING DATE: 1999-04-30  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: PatentIn Ver. 2.1

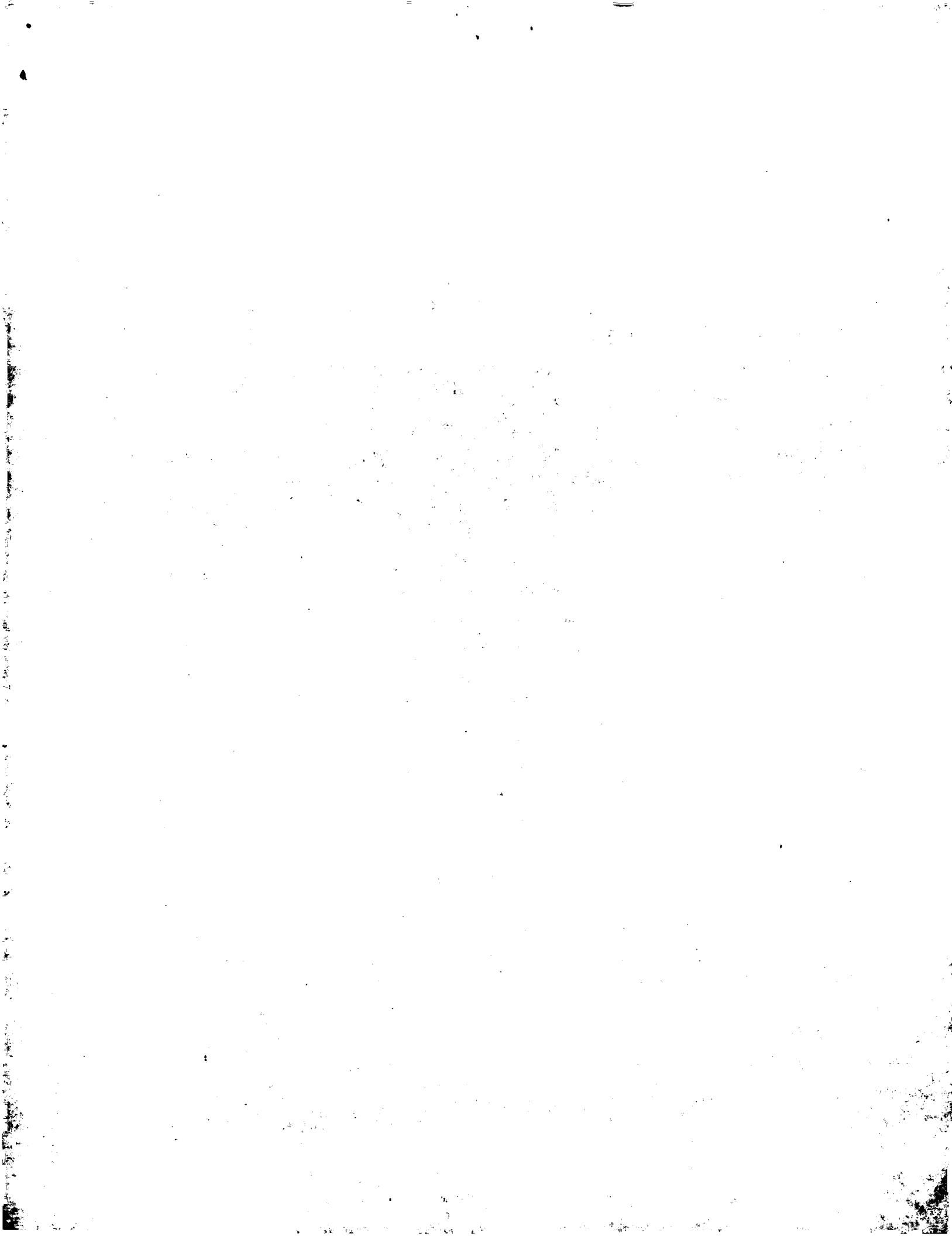
```
; SEQ ID NO 7
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-303-120B-7
```

```
RESULT 15
US-09-820-576-7 Application US/09820576
; Sequence 7, Application US/09820576
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Sutton, Christopher L.
; TITLE OF INVENTION: TBD-Associated Microbial Nucleic Acid
; FILE REFERENCE: P-PM 4646
; CURRENT APPLICATION NUMBER: US/09/820-576
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/303, 120
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7
LENGTH: 200
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-820-576-7
```

Query Match 19.2%; Score 94.5; DB 17; Length 200;  
 Best Local Similarity 26.9%; Pred. No. 0.013; Mismatches 34; Indels 3; Gaps 1;  
 Matches 21; Conservative 20; Mismatches 34; Indels 3; Gaps 1;

```
Qy 1 DLASAVGIQSIFHFKSKDELRAVMEETIHYNTAMMRASLEAST'RVRLIRCE 60
Db 35 DIAGAGILGSLSLYHFAKKEEMDRLRGFLDMLFARYRDIVSTANPLERLQCLFMAS 94
Qy 61 LQSIMGGSGEAMAVALYE 78
Db 95 FEATIEHHHAQ---WVYQ 109
```

Search completed: March 4, 2002, 20:27:16  
 Job time: 184 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 4, 2002, 20:24:37 ; Search time 19.94 Seconds  
(without alignments)  
287.207 Million cell updates/sec

Title: US-09-966-608-2  
Perfect score: 493  
Sequence: 1 DLASAVGIQSGSIFHFKSK . . . . . SLSAEGQAHVLALRDVYEQI 100  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 261336 seqs, 57268997 residues

Total number of hits satisfying chosen parameters: 261336

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

## Database :

```
1: /cgns2_6/ptodata/1/paa/pctt_NEW_COMBO.pep:*
2: /cgns2_6/ptodata/1/paa/US09_608_6_NEW_COMBO.pep:*
3: /cgns2_6/ptodata/1/paa/US09_NEW_COMBO.pep:*
4: /cgns2_6/ptodata/1/paa/US09_NEW_COMBO.pep:*
5: /cgns2_6/ptodata/1/paa/US10_NEW_COMBO.pep:*
6: /cgns2_6/ptodata/1/paa/US10_NEW_COMBO.pep:*
7: /cgns2_6/ptodata/1/paa/US10_NEW_COMBO.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## Result No.

## Score

## Query Match Length

## DB ID

## Description

## Sequence

ALIGNMENTS									
RESULT	1	US-09-966-608-2							
; Sequence 2, Application US/0966608									
; GENERAL INFORMATION:									
; APPLICANT: Braun, Jonathan									
; ATTORNEY: Sutton, Christopher L.									
; TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid									
; TITLE OF INVENTION: Molecules									
; FILE REFERENCE: P-PM 4966									
; CURRENT APPLICATION NUMBER: US/09/966, 608									
; CURRENT FILING DATE: 2001-09-27									
; PRIORITY NUMBER: US 09/7303, 120									
; PRIORITY FILING DATE: 1999-04-30									
; PRIORITY APPLICATION NUMBER: US 09/820, 576									
; NUMBER OF SEQ ID NOS: 10									
; SOFTWARE: FASTSEQ for Windows Version 4.0									
; SEQ ID NO: 2									
; LENGTH: 100									
; TYPE: PT									
; ORGANISM: Unknown									
; FEATURE: OTHER INFORMATION: Microbial organism from the human gut:									
; US-09-966-608-2									
; QUERY MATCH Best Local Similarity 100.0%; Score 493; DB 5; Length 100;									
; MATCHES 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Query	Match	Oy	1	DLASAVGIQSGSIFHFKSKDEILRAVMEETHYNTAMMASL EASTVIRVALIRCE	60				
; 1 DLASAVGIQSGSIFHFKSKDEILRAVMEETHYNTAMMASL EASTVIRVALIRCE 60									
Db	1	DLASAVGIQSGSIFHFKSKDEILRAVMEETHYNTAMMASL EASTVIRVALIRCE	60						
; 1 DLASAVGIQSGSIFHFKSKDEILRAVMEETHYNTAMMASL EASTVIRVALIRCE 60									
Qy	61	LOSIMGGSGEAMAVLYEWRSLSAEGQAHVLALRDVYEQI	100						
; 1 LOSIMGGSGEAMAVLYEWRSLSAEGQAHVLALRDVYEQI 100									
Db	61	LOSIMGGSGEAMAVLYEWRSLSAEGQAHVLALRDVYEQI	100						
; 1 LOSIMGGSGEAMAVLYEWRSLSAEGQAHVLALRDVYEQI 100									
; RESULT 2									
; US-09-976-451-5									
; GENERAL INFORMATION:									
; APPLICANT: Braun, Jonathan									
; APPLICANT: Wei, Bo									
; APPLICANT: Forbes, Ashley									
; TITLE OF INVENTION: Methods of Diagnosing and Treating									
; TITLE OF INVENTION: Crohn's Disease Using Pseudomonas Antigens									
; FILE REFERENCE: P-PM 4968									



APPLICANT: zelder, oskar  
 APPLICANT: habermann, gregor  
 TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL,  
 TITLE OF INVENTION: PROTEINS  
 FILE REFERENCE: BGI-1290P  
 CURRENT APPLICATION NUMBER: US/09/605,703B  
 CURRENT FILING DATE: 2000-06-27  
 PRIORITY APPLICATION NUMBER: 60/142,764  
 PRIORITY FILING DATE: 1999-07-08  
 PRIORITY APPLICATION NUMBER: 60/152,318  
 PRIORITY FILING DATE: 1999-09-03  
 NUMBER OF SEQ ID NOS: 2934  
 SEQ ID NO 2830  
 LENGTH: 218  
 TYPE: PRT  
 ORGANISM: Corynebacterium glutamicum  
 US-09-605-703B-2830

Query Match 18.3%; Score 90; DB 5; Length 218;  
 Best Local Similarity 29.4%; Pred. No. 0.016; Gaps 2;  
 Matches 20; Conservative 20; Mismatches 20; Indels 8; Gaps 2;

Qy 2 LASAVGIOGSFIFHFKSKDEIRRAVMEITHNTAMMRASLEASIV---RERVLALL 57  
 Db 41 IADAVGIRQASLYHFPSTKEFLTLKSTVERSTVL---AEDLSPLAGPEMRLWAI 96

Qy 58 RCELQSTM 65  
 Db 97 ASEVRLL 104

**RESULT** 7

US-09-966-608-B  
 Sequence 8, Application US/09966608  
 GENERAL INFORMATION:  
 APPLICANT: Braun, Jonathan  
 APPLICANT: Sutton, Christopher L.  
 TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid  
 TITLE OF INVENTION: Molecular  
 FILE REFERENCE: P-PM 4966  
 CURRENT APPLICATION NUMBER: US/09/966,608  
 CURRENT FILING DATE: 2001-03-27  
 PRIORITY APPLICATION NUMBER: US 09/303,120  
 PRIORITY FILING DATE: 1999-04-30  
 PRIORITY APPLICATION NUMBER: US 09/820,576  
 PRIORITY FILING DATE: 2001-03-28  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 8  
 LENGTH: 192  
 TYPE: PRT  
 ORGANISM: Aquifex aeolicus  
 US-09-966-608-B

Query Match 18.1%; Score 89; DB 4; Length 576775;  
 Best Local Similarity 41.3%; Pred. No. 7.6e+02; Gaps 4;  
 Matches 19; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

Qy 1 DLASAVGIOGSFIFHFKSKDEIRRAVMEITHNTAMMRASLEA 46  
 Db 365113 DIKEVGITEGAIRYRHTSKEEIKLMSLESI---TKELRHKEVA 365154

**RESULT** 9

US-08-895-611D-2  
 Sequence 2, Application US/08895611D  
 GENERAL INFORMATION:  
 APPLICANT: DIVERSA CORPORATION  
 APPLICANT: Swanson, Ronald V.  
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF THE AQUIFEX  
 TITLE OF INVENTION: AEOLOCUS GENOME, FRAGMENTS THEREOF, AND USES THEREOF  
 FILE REFERENCE: DIVER330  
 CURRENT APPLICATION NUMBER: US/08/895,611  
 CURRENT FILING DATE: 1997-07-16  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 576775  
 TYPE: PRT  
 ORGANISM: Aquifex aeolicus  
 US-08-895-611D-2

Query Match 18.1%; Score 89; DB 4; Length 576775;  
 Best Local Similarity 41.3%; Pred. No. 7.6e+02; Gaps 4;  
 Matches 19; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

Qy 1 DLASAVGIOGSFIFHFKSKDEIRRAVMEITHNTAMMRASLEA 46  
 Db 365113 DIKEVGITEGAIRYRHTSKEEIKLMSLESI---TKELRHKEVA 365154

**RESULT** 10

US-09-895-611D-2  
 Sequence 2, Application US/09895611D  
 GENERAL INFORMATION:  
 APPLICANT: DIVERSA CORPORATION  
 APPLICANT: Swanson, Ronald V.  
 TITLE OF INVENTION: Short, Jay M.  
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF THE AQUIFEX  
 TITLE OF INVENTION: AEOLOCUS GENOME, FRAGMENTS THEREOF, AND USES THEREOF  
 FILE REFERENCE: DIVER330  
 CURRENT APPLICATION NUMBER: US/09/895,611D  
 CURRENT FILING DATE: 1997-07-16  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 576775  
 TYPE: PRT  
 ORGANISM: Aquifex aeolicus  
 US-09-895-611D-2

Query Match 18.1%; Score 89; DB 4; Length 576775;  
 Best Local Similarity 41.3%; Pred. No. 7.6e+02; Gaps 4;  
 Matches 19; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

Qy 1 DLASAVGIOGSFIFHFKSKDEIRRAVMEITHNTAMMRASLEA 46  
 Db 365113 DIKEVGITEGAIRYRHTSKEEIKLMSLESI---TKELRHKEVA 365154

**RESULT** 8

US-08-895-611-2  
 Sequence 2, Application US/08895611  
 GENERAL INFORMATION:  
 APPLICANT: DIVERSA CORPORATION  
 APPLICANT: Swanson, Ronald V.  
 APPLICANT: Short, Jay M.  
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF THE AQUIFEX

Query Match 18.1%; Score 89; DB 5; Length 576775;  
 Best Local Similarity 41.3%; Pred. No. 7.6e+02;  
 Matches 19; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

Qy 1 DLASAVGQIQSFSFHHFSKDEILRAVMEETIHNTAMMRASLEA 46  
 Db 365113 DIAREVGITEGAIRYRHFTSKEEIKSILLESI---TKELRHKLEVA 365154

RESULT 11  
 US-015-127-12199  
 ; Sequence 12199, Application US/10015127  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bower, Stanley G.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof  
 ; FILE REFERENCE: 38-10(15606)B  
 ; CURRENT APPLICATION NUMBER: US/10/015,127  
 ; CURRENT FILING DATE: 2001-10-29  
 ; PRIOR APPLICATION NUMBER: US 60/252,455  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 14357  
 ; SEQ ID NO 12199  
 ; LENGTH: 189  
 ; TYPE: PRT  
 ; ORGANISM: Sphingomonas elodea  
 ; US-10-015-127-12199

RESULT 12  
 US-09-897-516-4814  
 ; Sequence 4814, Application US/09897516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51047)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215, 161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8409  
 ; SEQ ID NO 4814  
 ; LENGTH: 196  
 ; TYPE: PRT  
 ; ORGANISM: Xenorhabdus sp.  
 ; US-09-897-516-4814

Query Match 17.4%; Score 86; DB 6; Length 189;  
 Best Local Similarity 35.2%; Pred. No. 0.037; Mismatches 21; Indels 6; Gaps 2;

Matches 21; Conservative 13; Mismatches 18; Indels 6; Gaps 2;

\*Qy 2 LASAVGQIQSFSFHHFSKDEILRAVMEETIHNTAMMRASLEA--TVRERVAL 56  
 Db 14 LAQAVGVSPGAPYHFKDRALLAI--ALHGEALGAAAHASSNATAKDLLA 68

RESULT 13  
 US-09-897-516-5762  
 ; Sequence 5762, Application US/09897516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51047)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215, 161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8409  
 ; SEQ ID NO 5762  
 ; LENGTH: 183  
 ; TYPE: PRT  
 ; ORGANISM: Xenorhabdus sp.  
 ; US-09-897-516-5762

Query Match 16.1%; Score 79.5; DB 5; Length 183;  
 Best Local Similarity 27.4%; Pred. No. 0.19; Mismatches 31; Indels 21; Gaps 6;

Matches 31; Conservative 23; Mismatches 38; Indels 21; Gaps 6;

Qy 1 DLASAVGQIQSFSFHHFSKDEILRAVMEETIHNTAMMRASLEASTVRVAL--AII 57  
 Db 40 EVAREAGIAPTFSYRFHRDVDEGLTWDE---SGMLRQMRQA--RORIAKGSII 92

Qy 58 RCEQSIM--GGSGEAMVALVYEWRSLSAEGOA-----HVLA-LRDVEQ 99  
 Db 93 RISVSYTMEFIGHNPNAFRLLRRERSGSAEFAVAREIQHFIELADYLEQ 145

RESULT 14  
 US-09-897-516-5915  
 ; Sequence 5915, Application US/09897516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51047)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215, 161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8409  
 ; SEQ ID NO 5915  
 ; LENGTH: 193  
 ; TYPE: PRT  
 ; ORGANISM: Xenorhabdus sp.  
 ; US-09-897-516-5915

Query Match 16.1%; Score 79.5; DB 5; Length 193;  
 Best Local Similarity 27.0%; Pred. No. 0.21; Mismatches 27; Indels 5; Gaps 3;

Matches 27; Conservative 18; Mismatches 50; Indels 5; Gaps 3;

Qy 2 LASAVGQIQSFSFHHFSKDEILRAVMEETIH-YNTAMMRASLEASTVRVALJRC 59  
 Db 35 IARKKAGVSNGLISHYFRKRNKLEAMRHLHOLOMAVARRLRLDNATPLHRLKATIEG 94

QY 60 ELQSINGGSEAMAVLYEVRSLRSEGQAHVLALRDVYEQ 99  
 US-09-815-242-5102  
 ; Sequence 5102, Application US/09815242  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in  
 PROKARYOTES  
 TITLE OF INVENTION: Prokaryotes  
 CURRENT APPLICATION NUMBER: US/09/815, 242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191, 078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206, 848  
 PRIOR FILING DATE: 2000-03-23  
 PRIOR APPLICATION NUMBER: 60/207, 727  
 PRIOR FILING DATE: 2000-03-26  
 PRIOR APPLICATION NUMBER: 60/242, 578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253, 625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257, 931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269, 308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEO ID NO 5102  
 LENGTH: 194

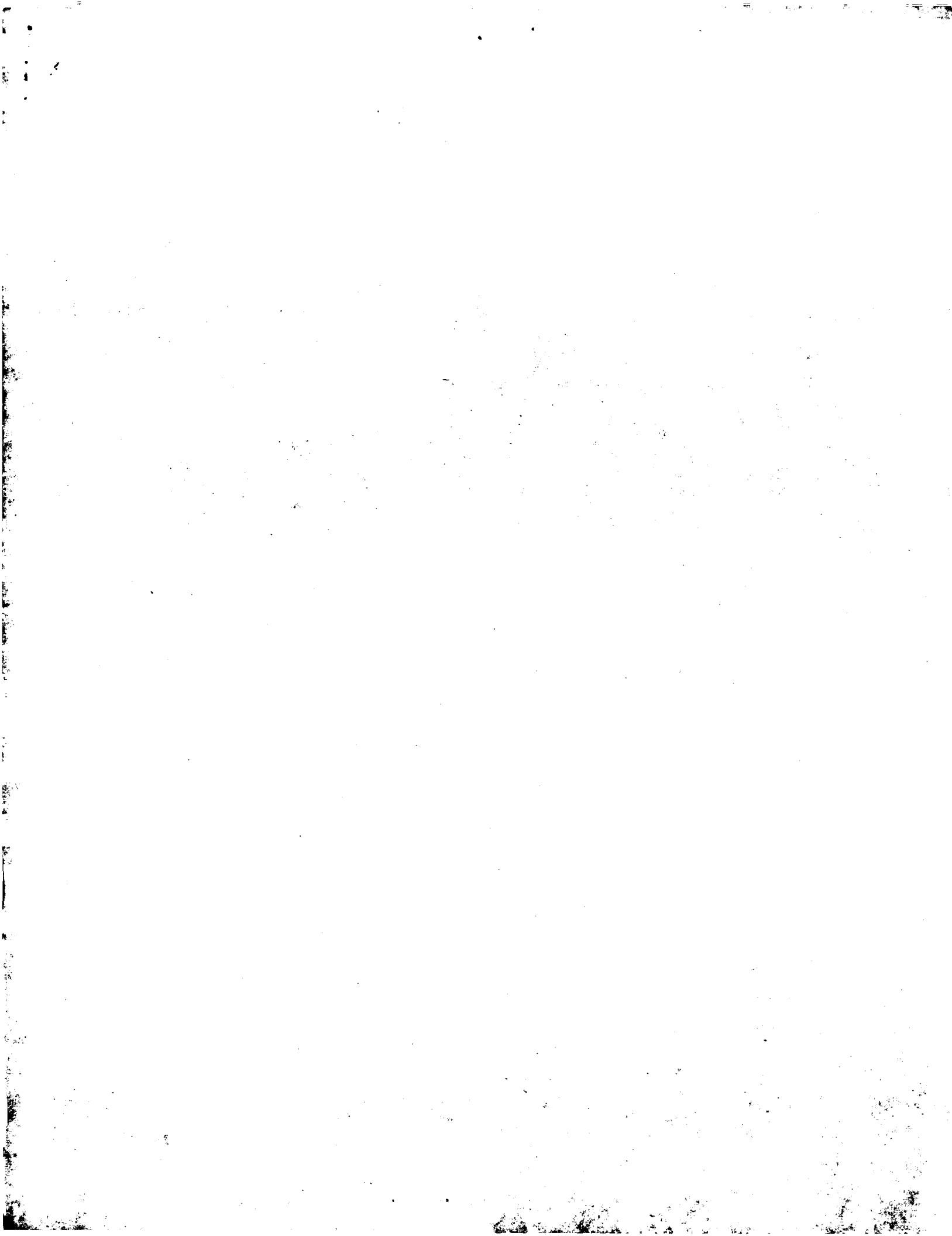
; ORGANISM: Pseudomonas aeruginosa

US-09-815-242-5102

Query Match Score 77.5; DB 5; Length 194;  
 Best Local Similarity 15.7%; Pred. No. 0.35; Mismatches 9; Gaps 3;  
 Matches 25; Conservative 23; Indels 9;

QY	1	DLASAVGIOGSISIHFHKSKDEILRAVMEETIH	YNTAMMRASLEASTYRERYLALIRC	59
Db	32	ELQSAGVPGKSFHYFKSKREQFGALLEDFRYVLADMDFRSAPGLNARERLMSWOK	91	
QY	60	ELQSINGGSEAMAVLYEVRSLRSEGQAHVLALRD	95	
Db	92	WLDNACPPCDEQRCLV---KLSAEVADLSESMRITLRD	127	

Search completed: March 4, 2002, 20:31:16  
 Job time: 399 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 4, 2002, 20:22:32 ; Search time 27.71 seconds

(without alignments)  
267.316 Million cell updates/sec

Title: US-09-966-608-2

Perfect score: 493

Sequence: DLASAVGIQSGSIFHHFKSK . . . . . SLSAEGQAHVLALRDVYEQI 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 7407290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101:\*

1: /SIDS2/gcadata/geneseq/geneseq/AAI980.DAT:\*

2: /SIDS2/gcadata/geneseq/geneseq/AAI981.DAT:\*

3: /SIDS2/gcadata/geneseq/geneseq/AAI982.DAT:\*

4: /SIDS2/gcadata/geneseq/geneseq/AAI983.DAT:\*

5: /SIDS2/gcadata/geneseq/geneseq/AAI984.DAT:\*

6: /SIDS2/gcadata/geneseq/geneseq/AAI985.DAT:\*

7: /SIDS2/gcadata/geneseq/geneseq/AAI987.DAT:\*

8: /SIDS2/gcadata/geneseq/geneseq/AAI988.DAT:\*

9: /SIDS2/gcadata/geneseq/geneseq/AAI989.DAT:\*

10: /SIDS2/gcadata/geneseq/geneseq/AAI990.DAT:\*

11: /SIDS2/gcadata/geneseq/geneseq/AAI991.DAT:\*

12: /SIDS2/gcadata/geneseq/geneseq/AAI992.DAT:\*

13: /SIDS2/gcadata/geneseq/geneseq/AAI993.DAT:\*

14: /SIDS2/gcadata/geneseq/geneseq/AAI994.DAT:\*

15: /SIDS2/gcadata/geneseq/geneseq/AAI995.DAT:\*

16: /SIDS2/gcadata/geneseq/geneseq/AAI996.DAT:\*

17: /SIDS2/gcadata/geneseq/geneseq/AAI997.DAT:\*

18: /SIDS2/gcadata/geneseq/geneseq/AAI998.DAT:\*

19: /SIDS2/gcadata/geneseq/geneseq/AAI999.DAT:\*

20: /SIDS2/gcadata/geneseq/geneseq/AAI2000.DAT:\*

21: /SIDS2/gcadata/geneseq/geneseq/AI2001.DAT:\*

22: /SIDS2/gcadata/geneseq/geneseq/AI2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	493	100.0	100	21 AAB19839 Inflammatory bowel
2	90	18.3	222	22 AAG90710 C glutamicum prote
3	83	18.3	214	22 AAU04043 Streptococcus coel
4	78.5	15.9	118	21 AAB40525 Human ORFX ORF89
5	73	14.8	3722	12 ARR10145 Cephalosporin anti
6	70.5	14.3	188	22 AAG91446 C glutamicum prote
7	69	14.0	80	21 ARR4577 Virulence gene pro
8	69	14.0	2818	13 ARR2268 Nf1 gene product.
9	68	13.8	509	19 AAW47533 Amino acid sequenc
10	68	13.8	509	19 AAW26792 Mouse TIE-2 receptor
11	13.8	509	21 AAY90398 Mouse TIE ligand-3	

12	67.5	13.7	508	17 AAR90296 Protein having pro
13	67.5	13.7	508	18 AAW1604 Arabidopsis thalia
14	67.5	13.7	508	18 Arabidopsis protoporphyrinogen oxidase
15	67.5	13.7	508	19 AAW5254 Arabidopsis thaliana
16	67.5	13.7	508	19 Arabidopsis thal
17	67.5	13.7	508	22 Arabidopsis protoporphyrinogen ABC
18	66.5	13.5	136	22 AAB62490 P. chrysogenum ABC
19	65.5	13.3	195	19 AAW42428 Glucuronide repres
20	65.5	13.3	411	19 AAW80938 Human kidney repre
21	65.5	13.3	414	19 AAW80941 Human kidney lecit
22	65.5	13.3	444	19 AAW80943 Human kidney lecit
23	64.5	13.1	195	21 AAG12675 Arabidopsis thalia
24	64.5	13.1	326	21 AAG1674 Arabidopsis thalia
25	64	13.0	156	21 AAB43399 Human ORF8X ORF303
26	64	13.0	783	22 AAG91263 C glutamicum prote
27	64	13.0	3768	12 AAR13753 Human neurofibroma
28	64	13.0	3778	12 AAR13895 Human polypeptide
29	63.5	12.9	1784	21 AAG07352 Arabidopsis thalia
30	63	12.8	2485	15 AAR5921 RAS associated GAP
31	63	12.8	2485	15 AAR5922 RAS associated GAP
32	63	12.8	2818	22 AAE05486 Human foetal prote
33	63	12.8	2818	22 AAM39070 C glutamicum prote
34	62.5	12.7	82	21 AAM40556 Human polypeptide
35	62.5	12.7	224	21 AAY81794 Arabidopsis thalia
36	62	12.6	76	22 AAM0641 Human polypeptide
37	62	12.6	217	22 AAG90819 Human polypeptide
38	62	12.6	219	12 AAB19432 Nucleotide sequenc
39	62	12.6	435	22 AAM39070 Human polypeptide
40	62	12.6	483	22 AAG51229 Arabidopsis thalia
41	62	12.6	778	21 AAG51228 Arabidopsis thalia
42	62	12.6	806	21 AAG51227 Arabidopsis thalia
43	62	12.6	927	21 AAG51227 Human polypeptide
44	61.5	76	14 AAR36783 Protein coded by 5	
45	61.5	12.5	211	22 AAG90850 C glutamicum prote

## ALIGNMENTS

RESULT 1  
ID AAB19839 standard; Protein: 100 AA.  
AC AAB19839;  
DT 05-MAR-2001 (first entry)

DE Inflammatory bowel disease associated antigen I-2.  
XX I-2; inflammatory bowel disease; IBD; Crohn's disease; therapy;  
KW diagnosis; vaccine.

OS Unidentified microorganism.  
XX WO200666067-A2.

PD 09-NOV-2000.  
XX PF 28-APR-2000; 2000WO-US11473.

PR 30-APR-1999; 99US-0303120.  
XX PA (RECG ) UNIV CALIFORNIA.

PI Braun J, Sutton C;  
XX DR WPI: 2000-687440-67.  
DR N-PSDB; AAR89029.

PT Inflammatory bowel disease (IBD) associated I-2 polypeptides useful for  
PT diagnosing IBD, vaccinating against IBD and for identifying agents for  
PT treating IBD -

CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Coryneform bacterium, and identifying a homologue of a gene derived  
 CC from coryneform bacterium. Coryneform bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described  
 CC in the exemplification of the invention.  
 Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.

Sequence	222 AA;
SQ	
Query Match	18.3%
Best Local Similarity	29.4%
ID	AAU04043
AC	AAU04043;
XX	
XX	23-OCT-2001 (first entry)
DT	
DE	Streptococcus coelicolor Mmfr protein.
XX	
KW	SCP1; methylenomycin cluster; mmrc; MmYR; Mmfp; Mmfl;
MmFL; Mmfr; MmYt; MmYO; MmYG; MmYJ; Mmfr; heterologous gene expression.	
XX	
OS	Streptococcus coelicolor.
OS	
PN	WO200148228-A1.
XX	
PD	05-JUL-2001.
XX	
PF	20-DEC-2000; 2000WO-GB04972.
XX	
PR	23-DBC-1999; 99GB-0030477.
XX	
PA	(PLAN-) PLANT BIOSCIENCE LTD.
XX	
PI	Charter KF, Bruton CJ, O'Rouke SJ, Wietzorek AW;
XX	
DR	WPI: 2001-425675/45.
DR	N-P5B2; AAS07627.
DR	AAU04045, AAU04046.
XX	
PT	Novel expression cassette for expressing a nucleic acid of interest, derived from the regulatory region of methylenomycin gene cluster of SCP1 plasmid of Streptomyces coelicolor A3(2) -
PT	Claim 25; Fig 8e; 14pp; English.
XX	
CC	The sequence represents the Mmfr protein encoded by the mmfr gene carried on the expression cassette present on plasmid SCP1. The expression cassette is the regulatory region of the methylenomycin cluster (mmrc) from Streptomyces coelicolor A3(2), which encodes the MmYR, Mmfp, Mmfl, Mmfr, MmYt, MmYJ and partial Mmfr polypeptides. The expression cassette is useful for expressing a nucleic acid of interest, substantially only when the host cell culture reaches high cell density at or close to the stationary phase of host cell culture. In particular
CC	



PS disclosure; Fig 13; 67pp; Japanese.

CC This protein is encoded by ORF1 of the 23666bp sequence  
CC isolated from L.lactamgenus and comprising the genes for the  
CC cephalosporin biosynthetic enzymes listed in the keywords. Plasmids  
CC containing at least one of ORF's 1-9 can be used to transform  
CC microbes, such as bacteria or yeast.  
CC See also AAQ10191-2.

XX SQ sequence 3722 AA;

Query Match 14.8%; Score 73; DB 12; Length 3722;  
Best Local Similarity 26.3%; Pred. No. 18; Gaps 1;  
Matches 20; Conservative 12; Mismatches 28; Indels 16; Gaps 1;

QY 5 AVGIOQSIFHFKSKBEILRAVMEETIHYNTAMMRASLEESTVRVRLALIRCELSI 64  
Db 901 atgqggmlygsankssasndayymlslyhyrraidpaamegdawlladqrkypslr----- 954

QY 65 MGGSGEAMAVLVYEW 80  
Db 955 -----lifewr 960

RESULT 6

AAG91446 ID AAG91446 standard; Protein: 188 AA.

XX AC AAG91446; AC

CC particularly L-lysine. The present sequence is a protein described  
CC in the exemplification of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.  
XX SQ sequence 188 AA;

Query Match 14.3%; Score 70.5; DB 22; Length 188;  
Best Local Similarity 25.8%; Pred. No. 0.66; Gaps 9; Mismatches 23; Indels 37; Gaps 4;  
Matches 24; Conservative 25.8%; Pred. No. 0.66; Gaps 9; Mismatches 23; Indels 37; Gaps 4;

QY 2 LASAVGIGSGSTFHFRKSKEILRAV-----MEETIHYN-TAMMRASLEESTVRER 52  
Db 37 Ieattgksrgafhhgdkenflalaedarmaevsenglvavmrgnledp----er 92

QY 53 VIALRICELOSTIMGGSGEAMAVLVYWRSLSAE 85  
Db 93 -----ydwmsvrl 101

RESULT 7

ID AAB44577 standard; Protein: 80 AA.

XX AC AAB44577; AC

XX DT 08-FEB-2001 (first entry)

XX DE Virulence gene protein #57.

XX KW Virulence gene; antibacterial; vaccine; bacterial infection;  
septicemia; bronchopneumonia; rhinitis; wound infection.

XX OS Actinobacillus pleuropneumoniae.

XX PN WO20061724-A2.

XX PD 19-OCT-2000.

XX PP 06-KPR-2000; 2000WO-US09218.

XX PR 09-APR-1999; 990US-0128689.

PR 10-SEP-1999; 990US-0153453.

XX PA (PHAA ) PHARMACIA & UPJOHN INC.

XX PI Lowery DE, Fuller TE, Kennedy MJ;

XX DR WPI; 2000-647422/62.

DR N-PSB; AAC79652.

XX PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence  
PT genes, useful as a live attenuated vaccine against bacterial infections

PT

XX PS Claim 39; Page 288; 322pp; English.

XX CC The family Pasteurellaceae encompasses several pathogens that infect a  
CC wide variety of animals. The present invention relates to virulence genes  
CC from Pasteurellaceae. The present sequence is a protein encoded by one  
CC such virulence gene. The virulence genes of the present invention may be  
CC mutated in order to produce an inactive gene. The inactive virulence gene  
CC may in turn be used to produce a vaccine, which is useful for treating  
CC bacterial infections such as septicemias, bronchopneumonias, rhinitis and  
CC wound infections.

XX SQ Sequence 80 AA;

Query Match 14.0%; Score 69; DB 21; Length 80;  
Best Local Similarity 41.4%; Pred. No. 0.32;

Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 2 LASAVGIOSGSIFHHFKSKDEILRAVNEE 30  
 ID AAR22268 standard; Protein; 2818 AA.  
 Db 42 iaseaglatgtiyrtktdelldclaeq 70  
 XX  
 AC AAR22268;  
 XX  
 DT 06-MAY-1992 (first entry)  
 XX  
 DE Nf1 gene product.  
 XX  
 KW von Recklinghausen neurofibromatosis disease; autosomal dominant; gene therapy.  
 KW  
 OS Homo sapiens.  
 PN W09200387-A.  
 XX  
 PD 09-JAN-1992.  
 PF 28-JUN-1991; 91WO-US04624.  
 XX  
 PR 29-JUN-1990; 90US-0547090.  
 XX  
 PA (UNMI ) UNIV OF MICHIGAN.  
 PI Collins FS, Wallace MR, Marchuk DA, Andersen LB, Gutmann DH;  
 XX  
 DR WPI; 1992-041568/05.  
 DR-N-PSDB; AAQ20602.  
 XX  
 PT DNA sequences to von-Recklinghausen neurofibromatosis gene - and derived amino acid sequences and probes for screening Nf1 in early stages of disease  
 PT  
 XX  
 PS Claim 25; Page 67; 12pp; English.

XX  
 CC This is the amino acid sequence of the von Recklinghausen neurofibromatosis (Nf1) gene product. It and antibodies raised to it can be used in hybridisation and immunological assays to screen for the presence of a normal or defective Nf1 gene product. Functional assays to measure levels of gene function can also be used for diagnosis or to monitor treatment. Patient therapy through supplementation with the normal Nf1 product which can be produced by recombinant techniques is also possible.  
 CC  
 XX  
 SQ Sequence 2818 AA;

Query Match 14.0%; Score 69; DB 13; Length 2818;  
 Best Local Similarity 29.1%; Pred. No. 40;  
 Matches 25; Conservative 15; Mismatches 28; Indels 18; Gaps 4;

QY 12 SIFFHFKSKDEILRAVNEE-----YNTAMMRASLEE-----ASTVRRVRLALTRCEL 61  
 ID 708 sefrhiceeadircavdevshmlpnyntfmefasvsnmnmstgraalqrvmallr-ri 766  
 AC QY 62 OSMGGSGEAMAVLVVWRSLSAEG 87  
 DB 767 eiptagteaa-----wedthakwq 785  
 XX  
 DE Mouse TIE-2 receptor ligand-3.  
 XX  
 KW TIE-2 receptor ligand; TIE ligand-3; tyrosine kinase; mouse;  
 KW angiogenesis; vascularisation; blood vessel growth; tumour;  
 KW therapy; receptorbody.  
 XX  
 OS Mus musculus.

RESULT 8  
 AAR22268  
 ID AAR22268 standard; Protein; 2818 AA.  
 AC AAR22268;  
 XX  
 DT 01-AUG-1997; 97WO-US13557.  
 XX  
 PR 25-OCT-1996; 96US-0740223.  
 PR 02-AUG-1996; 96US-0022499.  
 XX  
 PA (REGE-) REGENERON PHARM INC.  
 XX  
 PI Davis S, Yancopoulos GD;  
 XX  
 DR N-PSDB; AAIV18620.  
 XX  
 PR Modified human TIE-2 receptor ligand(s) - useful for promoting wound healing.  
 XX  
 PS Example 20; Fig 21; 202pp; English.  
 XX  
 CC This is the amino acid sequence of the murine TIE ligand 3, used in the method of the invention, involving the production of TIE-2 ligands which promote healing. The nucleic acids, vectors and host cell used in the method of the invention are useful for the recombinant production of the ligands. The ligands, etc. are useful for blocking blood vessel growth, promoting neovascularisation, promoting the growth or differentiation of a cell expressing the TIE receptor, blocking the growth or differentiation of a cell expressing the TIE receptor and for attenuating or preventing tumour growth in a human.  
 CC  
 XX  
 SQ Sequence 509 AA;

Query Match 13.8%; Score 68; DB 19; Length 509;  
 Best Local Similarity 28.4%; Pred. No. 5 2;  
 Matches 21; Conservative 18; Mismatches 29; Indels 6; Gaps 2;

QY 27 VNEETHYNTAMMRASLEE-----ASTVRRVRLALTRCEL 61  
 ID 165 vinqtlumtkmlensl---stnklqrqmlnsrelqrqgn---raletrlqaleaqn 218  
 AC QY 87 QAHVALRDVQEIQI 100  
 DB 219 qaqqlslqekreql 232  
 RESULT 10  
 AAW26792  
 ID AAW26792 standard; Protein; 509 AA.  
 AC AAW26792;  
 XX  
 DT 22-JUN-1998 (first entry)  
 XX  
 DE Mouse TIE-2 receptor ligand-3.  
 XX  
 KW TIE-2 receptor ligand; TIE ligand-3; tyrosine kinase; mouse;  
 KW angiogenesis; vascularisation; blood vessel growth; tumour;  
 KW therapy; receptorbody.  
 XX  
 OS Mus musculus.

RESULT 9  
 AAW47533  
 ID AAW47533 standard; Protein; 509 AA.  
 XX  
 AC AAW47533;



PA (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Volrath S, Ward ER;  
 PT  
 PT herbicide-resistant enzymes, and plants containing such constructs  
 XX  
 DR WPI; 1996-049687/05.  
 DR N-PSDB; AAT11674.  
 XX  
 PT DNA encoding eukaryotic proto:porphyrinogen oxidase and herbicide  
 PT resistant mutants - used to make herbicide resistant plants and for  
 diagnosis and treatment of variegate porphyria  
 XX  
 PS Claim 6; Page 77-79; 118pp; English.  
 XX  
 CC DNA encoding a modified protoporphyrinogen oxidase (PPO) or a  
 CC chimeric gene comprising a promoter (pref. active in a plant) linked  
 CC to such a DNA or DNA encoding a wild type PPO can be used (1) to  
 CC impart herbicide resistance to plants; (2) for treating and  
 CC diagnosing deficient PPO activity in animals (esp. variegate  
 CC porphyria); and (3) for the production of recombinant PPO which is  
 CC useful as an assay reagent and in rational design of new inhibitory  
 CC herbicides. Herbicide resistant PPO genes can also be used to  
 CC select plants transformed with a transgene and probes derived from  
 CC the genes can be used to quantify levels of PPO mRNA.  
 XX  
 SQ Sequence 508 AA;  
 Query Match 13.7%; Score 67.5; DB 17; Length 508;  
 Best Local Similarity 26.7%; Pred. No. 6; Mismatches 20; Conservative 16; Indels 11; Gaps 3; Matches 20; Mismatches 28; Gaps 3;  
 QY 15 HRFKSKDEILRAVM---EETHNTAMMRAS---LEEASTVVERVALIQLCELOSIM 65  
 Db 366 hgfkltglftfssmmfpdrpsdvhlyttffiggrnqelakast--deikqvvtstdqlqrll 423  
 QY 66 GGSGEAMAWLVYEWIR 80  
 Db 424 gvegepvsvhyywr 438  
 RESULT 13  
 AAM41604 ID AAM41604 standard; Protein: 508 AA.  
 AC AAM41604;  
 XX DT 20-APR-1998 (first entry)  
 DE Arabidopsis thaliana protox-2.  
 XX KW Protoporphyrinogen oxidase-2; protox-2; promoter; gene isolation;  
 KW herbicide resistance; breeding programme; probe; gene isolation;  
 genomic mapping.  
 XX OS Arabidopsis thaliana.  
 PN WO9732028-A1.  
 PD 04-SEP-1997.  
 FF 27-FEB-1997; 97WO-US03313.  
 PR 21-JUN-1996; 96US-0020003.  
 PR 28-FEB-1996; 96US-0012705.  
 PR 28-FEB-1996; 96US-0013612.  
 XX PA (NOVS ) NOVARTIS AG.  
 XX PI Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;  
 PT DR WPI; 1997-44863/41.  
 DR N-PSDB; AAT86130.  
 XX PT New DNA encoding plant protoporphyrinogen oxidase enzyme - and  
 PT herbicide resistant mutants, useful to prepare plants resistant to  
 PT herbicide which therefore kills undesired vegetation only  
 XX PS Claim 95; Page 102-105; 196pp; English.  
 XX CC This protein comprises Arabidopsis protoporphyrinogen oxidase  
 CC (protox-2), an enzyme that catalyses the oxidation of  
 CC protoporphyrinogen IX to protoporphyrin IX. Its amino acid sequence  
 XX

CC was deduced from a previously obtained *Protox-1* cDNA clone (see AAT86130). *Arabidopsis protox-2* can be modified to render it resistant to protox inhibitors and hence tolerant of certain herbicides. Plants, especially crop plants, may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form, or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme, such as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed rape, rice and sorghum (see AAW25738-48). Application of herbicide will then kill undesired vegetation only. Protox enzymes can also be expressed in transformed host cells and used to identify inhibitors of protox enzyme activity, i.e. herbicide candidates, or to design herbicide tolerant forms of the enzyme.

XX sequence 508 AA;  
SQ

Query Match 13.7%; Score 67.5; DB 18; Length 508;  
Best Local Similarity 26.7%; Pred. No. 6;  
Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

Qy 15 HHEKSKDEILRAVM----BETIINYNTAMRAS---LEEASTVRRVALIRCELOSIM 65  
Db 366 hgfkltglfssmmfpdrspdsvhlyttffggsrqelakast--deikqvvtsdlqrll 423

Qy 66 GGSGEAMAVLYYENR 80  
Db 424 gvegepvsvhyywr 438

RESULT 15

AAW51254 ID AAW51254 standard; Protein: 508 AA.  
XX AC AAW51254;  
XX

DT 14-AUG-1998 (first entry)  
DE *Arabidopsis* proto-porphyrinogen oxidase-2 (*protox-2*).  
XX KW Proto-porphyrinogen oxidase; protox; herbicide tolerant; resistance;  
KW inhibitor; variegate porphyria.  
XX OS *Arabidopsis thaliana*.  
XX PN US5767373.A.  
XX PD 16-JUN-1998.  
XX PP 16-JUN-1994; 94US-0261198.  
XX PR 06-JUN-1995; 95US-0472028.  
PR 16-JUN-1994; 94US-0261198.  
XX PA (NOVS ) NOVARTIS FINANCE CORP.  
XX PI Volrath S, Ward ER;  
XX DR WPI; 1998-361821/31.  
DR N-PSDB; AAV07252.

XX DNA encoding eukaryote herbicide resistant proto-porphyrinogen oxidase - useful for producing recombinant plants having functional enzyme, to be grown in the presence of herbicides  
XX PS Claim 6; Columns 21-24; 43pp; English.  
XX  
PT The invention relates to eukaryotic DNA sequences coding for native proto-porphyrinogen oxidase (protox) or modified forms of the enzyme which are herbicide tolerant. Plants having altered protox activity may be bred or engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form or through

CC increased levels of expression of the native protox gene, or they may be transformed with modified eukaryotic or prokaryotic protox coding sequences or wild type prokaryotic protox sequences which are herbicide tolerant. In the human condition variegate porphyria, an autosomal dominant disorder characterised by neuropsychiatric symptoms, decreased protox activity is detected. Thus the DNA molecules can be used to provide probes to detect and quantify protox levels in the diagnosis of diseases associated with the enzyme. The present sequence represents *Arabidopsis* proto-porphyrinogen oxidase-2 (*protox-2*).

XX Sequence 508 AA;

Query Match 13.7%; Score 67.5; DB 19; Length 508;  
Best Local Similarity 26.7%; Pred. No. 6;  
Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

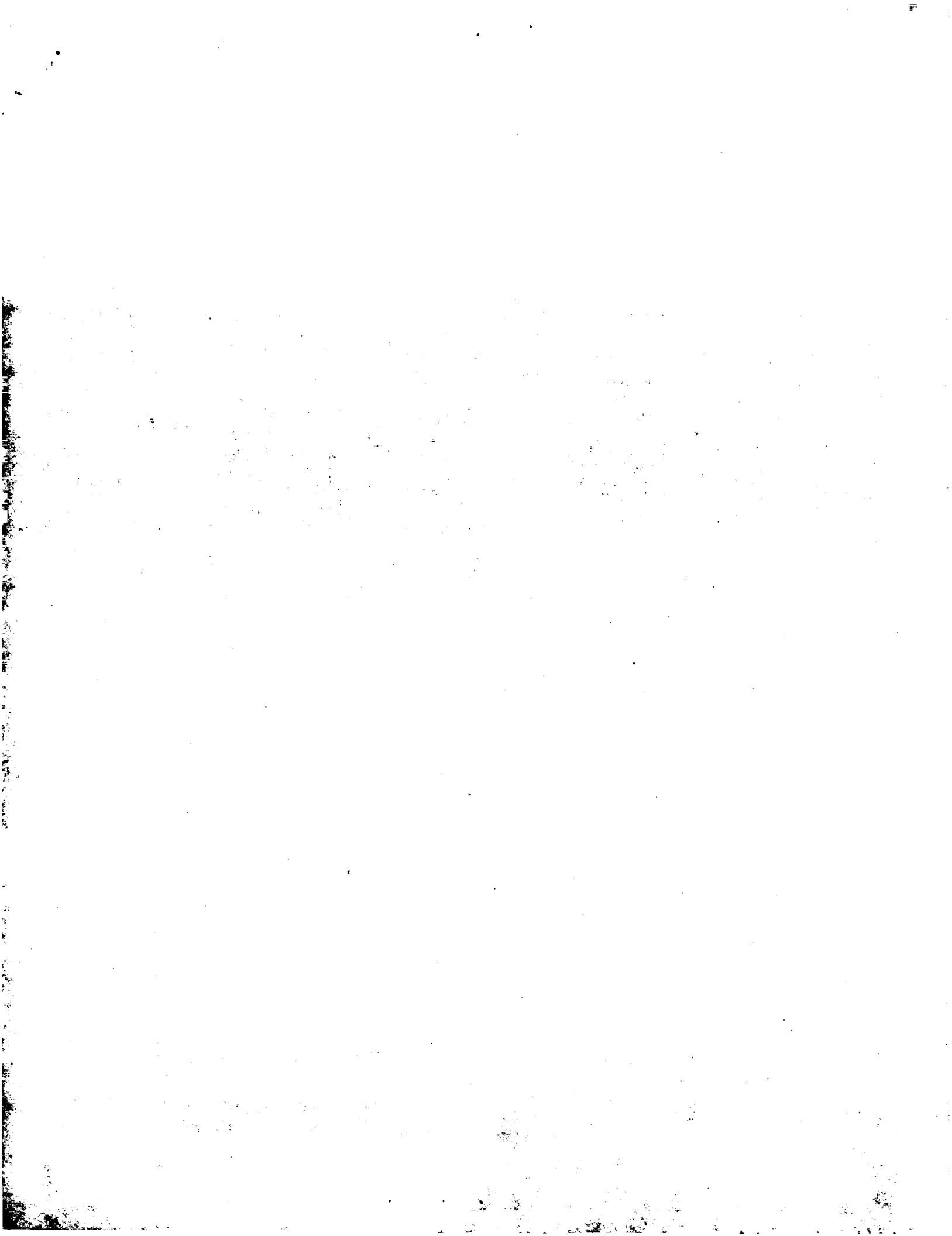
Qy 15 HHEKSKDEILRAVM----BETIINYNTAMRAS---LEEASTVRRVALIRCELOSIM 65  
Db 366 hgfkltglfssmmfpdrspdsvhlyttffggsrqelakast--deikqvvtsdlqrll 423

Qy 66 GGSGEAMAVLYYENR 80  
Db 424 gvegepvsvhyywr 438

Search completed: March 4, 2002, 20:24:06  
Job time: 94 sec

Tue Mar 5 07:57:15 2002

us-09-966-608-2.rag



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 4, 2002, 20:23:12 ; Search time 15.26 Seconds  
(without alignments)  
147.466 Million cell updates/sec

Title: US-09-966-608-2  
Perfect score: 493  
Sequence: 1 DLASAVGIQSGSIFHHFKSK . . . . . SLSAEGQAHVLALRDVYEQI 100  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%, Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/1aa/5\_A-COMB\_pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5\_B\_COMB\_pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6\_A\_COMB\_pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6\_B\_COMB\_pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB\_pep: \*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1\_pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	68	13.8	503	4 US-08-740-223A-11
2	68	13.8	509	2 US-08-655-926-8
3	68	13.8	509	4 US-08-740-223A-10
4	67.5	13.7	508	1 US-08-472-028A-4
5	67.5	13.7	508	2 US-08-931-4
6	67.5	13.7	508	3 US-08-930-323-4
7	67.5	13.7	508	3 US-09-050-603A-4
8	67.5	13.7	508	3 US-09-022-420B-4
9	67.5	13.7	508	4 US-09-071-266-4
10	67.5	13.7	508	4 US-09-196-266-4
11	67.5	13.7	508	4 US-09-015-683-4
12	67.5	13.7	508	4 US-09-191-938-4
13	67	13.6	2818	2 US-08-449-933-2
14	67	13.6	2818	4 US-07-066-09A-2
15	65.5	13.3	195	2 US-08-882-704A-2
16	64	13.0	768	2 US-08-222-617A-5
17	64	13.0	3666	2 US-08-222-617A-12
18	64	13.0	3727	2 US-08-222-617A-2
19	64	13.0	3778	2 US-08-222-617A-2
20	63.5	12.9	1784	4 US-09-040-730-2
21	63.5	12.9	1784	4 US-08-552-426A-2
22	63	12.8	2485	5 PCT-US94-0198-1
23	63	12.8	2485	5 PCT-US94-0198-2
24	63	12.8	2818	1 US-08-510-284-1
25	63	12.8	2818	1 US-08-411-389-2
26	63	12.8	2818	4 US-09-542-331-2
27	62	12.6	US-09-162-184-32	

## ALIGNMENTS

RESULT 1  
US-08-740-223A-11  
Sequence 11, Application US/08740223A  
Patent No. 6265564  
GENERAL INFORMATION:  
APPLICANT: Davis, et al.  
TITLE OF INVENTION: Expressed Ligand - Vascular  
TITLE OF INVENTION: Intercellular Signalling Molecule  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disquette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08740,223A  
FILING DATE: 25-OCT-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 60/022/999  
FILING DATE: 02-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 333  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 503 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: mtl3  
LOCATION: 1...503  
OTHER INFORMATION: mouse TIE ligand-3  
US-08-740-223A-11

Query Match 13.8%; Score 68; DB 4; Length 503;

Best local similarity 28.4%; Pred. No. 1; Matches 21; Conservative 18; Mismatches 29; Indels 6; Gaps 2; Result 2

Qy 27 VMEETIHTYNTAMMRSLEASTVRERVALIRELQSTMGSGEAMAVLYENRSLSAEG 86  
 Db 159 VLNQTLHMKTMQMLNSL--STNKLERQMLMQSRELQRQGRN---RALETRLQALEQH 212

Qy 87 QAHVLALRDVYEQT 100  
 Db 213 QAQLNSLQKREQL 226

RESULT 2

; Sequence 8, Application US/08665926

; Patent No. 5851797

; GENERAL INFORMATION:

; APPLICANT: Valenzuela et al.

; TITLE OF INVENTION: TIE LIGAND-3, METHODS OF MAKING AND USES

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

; STREET: 777 Old Saw Mill River Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10591-6707

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/665,926

; FILING DATE: 19-JUN-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Robert J. Cobert

; REGISTRATION NUMBER: 36,108

; REFERENCE DOCKET NUMBER: REG 333

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 914-345-4000

; TELEFAX: 914-345-7721

; INFORMATION FOR SEQ ID NO: 10:

; PRIORITY APPLICATION NUMBER: 60/022,999

; PRIORITY DATE: 02-AUG-1996

; PRIORITY FILING NUMBER: 911A-345-7721

; PRIORITY FILING DATE: 02-AUG-1996

; PRIORITY COUNTRY: USA

; PRIORITY INVENTION TITLE: Intercellular Signalling Molecule

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

; STREET: 777 Old Saw Mill Road

; CITY: Tarrytown

; STATE: NY USA

; ZIP: 10591

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/740,223A

; FILING DATE: 25-OCT-1996

; CLASSIFICATION: 536

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: USN 60/022,999

; PRIORITY FILING DATE: 02-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Cobert, Robert J.

; REGISTRATION NUMBER: 36,108

; REFERENCE DOCKET NUMBER: REG 333

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 914-345-4000

; TELEFAX: 914-345-7721

; INFORMATION FOR SEQ ID NO: 10:

; PRIORITY APPLICATION NUMBER: 60/022,999

; PRIORITY PRIORITY FILING NUMBER: 911A-345-7721

; PRIORITY PRIORITY FILING DATE: 02-AUG-1996

; PRIORITY PRIORITY COUNTRY: USA

; PRIORITY PRIORITY INVENTION TITLE: Intercellular Signalling Molecule

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-Geigy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NJ USA

; ZIP: 07032

Query Match 13.8%; Score 68; DB 2; Length 509;  
 Best Local Similarity 28.4%; Pred. No. 1;  
 Matches 21; Conservative 18; Mismatches 29; Indels 6; Gaps 2;

Qy 27 VMEETIHTYNTAMMRSLEASTVRERVALIRELQSTMGSGEAMAVLYENRSLSAEG 86  
 Db 165 VLNQTLHMKTMQMLNSL--STNKLERQMLMQSRELQRQGRN---RALETRLQALEQH 218

Qy 87 QAHVLALRDVYEQT 100  
 Db 219 QAQLNSLQKREQL 232

RESULT 3

US-08-740-223A-10

; Sequence 10, Application US/08740223A

; Patent No. 6265564

; GENERAL INFORMATION:

; APPLICANT: Davis, et al.

; TITLE OF INVENTION: Expressed Ligand - Vascular

Query Match 13.8%; Score 68; DB 2; Length 509;  
 Best Local Similarity 28.4%; Pred. No. 1;  
 Matches 21; Conservative 18; Mismatches 29; Indels 6; Gaps 2;

Qy 27 VMEETIHTYNTAMMRSLEASTVRERVALIRELQSTMGSGEAMAVLYENRSLSAEG 86  
 Db 165 VLNQTLHMKTMQMLNSL--STNKLERQMLMQSRELQRQGRN---RALETRLQALEQH 218

Qy 87 QAHVLALRDVYEQT 100  
 Db 219 QAQLNSLQKREQL 232

RESULT 4

US-08-472-028A-4

; Sequence 4, Application US/08472028A

; Patent No. 5767373

; GENERAL INFORMATION:

; APPLICANT: Ward, Eric R

; APPLICANT: Vollrath, Sandra

; TITLE OF INVENTION: Manipulation of Protoporphyrinogen

; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-Geigy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NJ USA

; ZIP: 07032

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/472,028A  
 FILING DATE: 21-JUN-1996  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James, Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1748/CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919) 541-8614  
 TELEFAX: (919) 541-8614  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 508 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ; US-08-472-028A-4

Query Match 13.7%; Score 67.5; DB 1; Length 508;  
 Best Local Similarity 26.7%; Pred. No. 1.2;  
 Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

QY 15 HHPKSKDEIRAVM---EETIHYNTAMMAS---LEPASTYRVERVALIRCLQSIM 65  
 | | | : .| : | | : | | | : | | | : | | | : | | | : | | | : | | |  
 | | | : .| : | | : | | | : | | | : | | | : | | | : | | | : | | |  
 Db 366 HGFKTIGTLSSMMPPDRSPSDVHLYTTFIGGSRNQELAKAST--DELKVQVTSSDLQRLL 423

Qy 66 GGSGEAMAVLYEW 80  
 | | | : | | | : | | |  
 Db 424 GVEGEPVSVNHYIWR 438

---

RESULT 5  
 US-08-808-931-4  
 ; Sequence 4, Application US/08808931  
 ; Patent No. 5939602

GENERAL INFORMATION:  
 APPLICANT: Voirath, Sandra  
 APPLICANT: Johnson, Marie  
 APPLICANT: Porter, Sharon  
 APPLICANT: Ward, Eric  
 APPLICANT: Heifetz, Peter  
 TITLE OF INVENTION: DNA Molecules Encoding Plant  
 TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 5939602artis Corporation  
 STREET: 520 White Plains Road, P.O. Box 2005  
 CITY: Tarrytown  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10591-9005

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,931  
 FILING DATE: 28-FEB-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/012,705  
 FILING DATE: 28-FEB-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/013,612  
 FILING DATE: 28-FEB-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/020,003  
 FILING DATE: 21-JUN-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC 1847

RESULT 6  
 US-08-808-323-4  
 ; Sequence 4, Application US/08808323  
 ; Patent No. 6018105

GENERAL INFORMATION:  
 APPLICANT: Johnson, Marie  
 APPLICANT: Voirath, Sandra  
 APPLICANT: Ward, Eric  
 TITLE OF INVENTION: Promoters from Plant  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6018105artis Corporation  
 STREET: 520 White Plains Road, P.O. Box 2005  
 CITY: Tarrytown  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10591-9005

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,323  
 FILING DATE:  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/012,705  
 FILING DATE: 28-FEB-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/013,612  
 FILING DATE: 28-FEB-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/020,003  
 FILING DATE: 21-JUN-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC 1846

TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (919) 541-8587  
 ; FAX: (919) 541-8689  
 ; INFORMATION FOR SEO ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 508 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-808-323-4.

Query Match 13.7%; Score 67.5; DB 3; Length 508;  
 Best Local Similarity 26.7%; Pred. No. 1.2; Mismatches 20;  
 Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

QY 15 HHEFKSKDEILRAVM----EETIHYNTAMMRS---LEEAStVREVLALIRCELOSIM 65  
 Db 366 HGFKTLGTLFSSNMFPDRSPSDVHLYTFIGGSRNQELAKAST--DELKQVVTSDLQRLL 423

QY 66 GGSEAMAVLVYER 80  
 Db 424 GVEGEPVSVNHYWWR 438

RESULT 7  
 US 09-050-603A-4  
 Sequence 4, Application US/09050603A  
 Patent No. 6023012  
 GENERAL INFORMATION:  
 APPLICANT: Volrath, Sandra  
 APPLICANT: Johnson, Marie  
 APPLICANT: Potter, Sharon  
 APPLICANT: Ward, Eric  
 APPLICANT: Heifetz, Peter  
 TITLE OF INVENTION: DNA Molecules Encoding Plant  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6023012artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 COMPUTER: IBM PC compatible  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patientin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/050, 603A  
 FILING DATE: 30-MAR-1998  
 CLASSIFICATION: 800  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 08/808, 931  
 FILING DATE: 28-FEB-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/012, 705  
 FILING DATE: 28-FEB-1996  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 60/013, 612  
 FILING DATE: 28-FEB-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/020, 003  
 FILING DATE: 21-JUN-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meiers, J. Timothy  
 REGISTRATION NUMBER: 38, 241  
 REFERENCE/DOCKET NUMBER: CGC 1847  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919) 541-8587  
 TELEFAX: (919) 541-8689

RESULT 8  
 US-09-102-420B-4  
 Sequence 4, Application US/09102420B  
 Patent No. 6084155  
 GENERAL INFORMATION:  
 APPLICANT: Volrath, Sandra  
 APPLICANT: Johnson, Marie  
 APPLICANT: Ward, Eric  
 APPLICANT: Heifetz, Peter  
 TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN  
 NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6084155artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patientin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/102, 420B  
 FILING DATE: 22-JUN-1998  
 CLASSIFICATION: 800  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 09/059, 164  
 FILING DATE: 13-APR-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 09/050, 603  
 FILING DATE: 30-MAR-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 60/126, 430  
 FILING DATE: 11-MAR-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 08/808, 931  
 FILING DATE: 28-FEB-1997  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 60/012, 705  
 FILING DATE: 28-FEB-1996  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 60/013, 612  
 FILING DATE: 28-FEB-1996  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 60/020, 003  
 FILING DATE: 21-JUN-1996  
 PRIORITY DATA:  
 APPLICATION NUMBER: US 08/472, 028

INFORMATION FOR SEO ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 508 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ; US-09-050-603A-4

Query Match 13.7%; Score 67.5; DB 3; Length 508;  
 Best Local Similarity 26.7%; Pred. No. 1.2; Mismatches 20;  
 Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

QY 15 HHEFKSKDEILRAVM----EETIHYNTAMMRS---LEEAStVREVLALIRCELOSIM 65  
 Db 366 HGFKTLGTLFSSNMFPDRSPSDVHLYTFIGGSRNQELAKAST--DELKQVVTSDLQRLL 423

QY 66 GGSEAMAVLVYER 80  
 Db 424 GVEGEPVSVNHYWWR 438

FILING DATE: 06-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC 1847/CIP3

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919) 541-8587  
 TELEFAX: (919) 541-8689  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 508 amino acids

TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

; US-09-102-420B-4

Query Match 13.7%; Score 67.5; DB 3; Length 508;

Best Local Similarity 26.7%; Pred. No. 1.2; Mismatches 3;  
 Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

Qy 15 HHFKSKDEILRAVM---EETIHNNTAMRAS---LEASTVRERVLALIRCLOSM 65  
 Db 366 HGFKTGTGTLSSMMPPDRSPSDVHLYTTFIGGSRSRQELAKAST--DELKVWTSQRLIL 423

Qy 66 GGSGEAMAVLYEWR 80  
 Db 424 GVEGEPVSVNHYWYR 438

RESULT 9

US-09-071-296-4

; Sequence 4, Application US/09071296

Patent No. 617245

GENERAL INFORMATION:

APPLICANT: Ward, Eric R

APPLICANT: Voirath, Sandra

APPLICANT: Voirath, Sandra

TITLE OF INVENTION: Manipulation of Protoporphyrinogen

TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,296

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1748/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8614

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 508 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

; US-09-071-296-4

RESULT 10

US-09-196-268-4

; Sequence 4, Application US/09196268

Patent No. 6282837

GENERAL INFORMATION:

APPLICANT: Ward, Eric R

APPLICANT: Voirath, Sandra

APPLICANT: Voirath, Sandra

TITLE OF INVENTION: Manipulation of Protoporphyrinogen

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/196,268

FILING DATE: 06-JUN-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/261,198

FILING DATE: 16-JUN-94

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1748/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8614

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 508 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

; US-09-196-268-4

Query Match 13.7%; Score 67.5; DB 4; Length 508;

Best Local Similarity 26.7%; Pred. No. 1.2; Mismatches 3;

Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;

Qy 15 HHFKSKDEILRAVM---EETIHNNTAMRAS---LEASTVRERVLALIRCLOSM 65  
 Db 366 HGFKTGTGTLSSMMPPDRSPSDVHLYTTFIGGSRSRQELAKAST--DELKVWTSQRLIL 423

Qy 66 GGSGEAMAVLYEWR 80  
 Db 424 GVEGEPVSVNHYWYR 438

US-09-071-296-4

RESULT 11  
 Sequence 4, Application US/09015683  
 Patent No. 6288306  
 GENERAL INFORMATION:  
 APPLICANT: Ward, Eric R  
 TITLE OF INVENTION: Manipulation of Protoporphyrinogen  
 TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ciba-Geigy Corporation  
 STREET: 7 Skyline Drive  
 CITY: Hawthorne  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/191,998  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1748/CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 508 amino acids  
 TYPE: amino acid  
 TOPLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-015-683-4

RESULT 12  
 Query Match Best Local Similarity 13.7%; Score 67.5; DB 4; Length 508; Matches 20; Conservative 16; Mismatches 28; Indels 11; Gaps 3;  
 Oy 15 HIFKSKDELRAVM---EETIHYNTAMMAS---LEAESTVRERVLALRCESIM 65  
 Db 366 HGFKTIGTLFESSMMFPDRSPSYHLYTFIGGSRNLAKAST--DELKQVWTSIQRL 423  
 Oy 66 GGSGEAMAVLVYEW 80  
 Db 424 GVEGEPVSVNHYWR 438

RESULT 13  
 US-08-449-933-2  
 Sequence 2, Application US/08449933  
 Patent No. 5539195  
 GENERAL INFORMATION:  
 APPLICANT: Collins, Francis S.  
 APPLICANT: Wallace, Margaret R.  
 APPLICANT: Marchuk, Douglas A.  
 APPLICANT: Anderson, Lone B.  
 APPLICANT: Gutmann, David H.  
 TITLE OF INVENTION: Neurofibromatosis Gene  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 75 Page Mill Road  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/449,933  
 FILING DATE: 25-MAY-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Konski, Antoinette F.  
 REGISTRATION NUMBER: 34,202  
 REFERENCE/DOCKET NUMBER: 20344-20553.10  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 434-0792  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2818 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 POSITION IN GENOME:  
 CHROMOSOME SEGMENT: 17q11.2  
 FEATURE:  
 NAME/KEY: Cleavage-site  
 LOCATION: group(583..586, 815..818, 2573..2575, 2810..2813)  
 OTHER INFORMATION: /note= "Potential CAMP-dependent phosphorylation site"  
 OTHER INFORMATION: protein kinase recognition sites"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 2549..2556  
 OTHER INFORMATION: /note= "Potential tyrosine phosphorylation site"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: group(1264, 1276, 1358, 1377, 1389, 1390, 1391,  
 1395, 1396, 1400, 1423, 1426, 1429, 1430)  
 OTHER INFORMATION: /note= "Invariant residues within most statistically significant regions of similarity among the GAP family of proteins"  
 OTHER INFORMATION: most statistically significant regions of similarity among the GAP family of proteins"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: group(1264..1270, 1345..1407, 1415..1430)  
 OTHER INFORMATION: /note= "Most statistically significant regions of similarity among the GAP family of proteins"  
 OTHER INFORMATION: proteins"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 496  
 OTHER INFORMATION: /note= "At variance with previously published sequence which shows an ATG methionine codon rather than an ATA isoleucine codon"  
 OTHER INFORMATION: than an ATA isoleucine codon"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 1183  
 OTHER INFORMATION: /note= "At variance with previously published sequence. Shows an CTG leucine codon rather than the previously published CTC"  
 OTHER INFORMATION: previously published CTC"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 1555  
 OTHER INFORMATION: /note= "At variance with previously published sequence. Lacks an extra CAT histidine codon after this residue"  
 OTHER INFORMATION: this residue"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: (271..2772)  
 OTHER INFORMATION: /note= "Position of an 18 amino acid insertion (SEQ ID NO:10) representing an alternatively spliced product"  
 OTHER INFORMATION: spliced product"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: (1370..1371)  
 OTHER INFORMATION: /note= "Position of a 21 amino acid insertion representing an alternatively spliced product"  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 1125..1537  
 OTHER INFORMATION: /note= "NFL catalytic domain"

FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 2446..2818  
 OTHER INFORMATION: /note= "Corresponding amino acids for the PstI-HindIII fragment designated pMAL. HF3A. X"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 65..371  
 OTHER INFORMATION: /note= "Corresponding amino acids for the HpaI-PstI fragment designated pMAL. HF3A. P"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 65..1240  
 OTHER INFORMATION: /note= "Corresponding amino acids for the HpaI-XbaI fragment designated pMAL. HF3A. X"  
 PUBLICATION INFORMATION:  
 AUTHORS: Wallace, M.R. et al.  
 TITLE: Type I Neurofibromatosis Gene: Correction  
 JOURNAL: Science  
 VOLUME: 250  
 ISSUE: 12/21/90  
 PAGES: 1749-  
 DATE: 12/21/1990  
 RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818  
 PUBLICATION INFORMATION:  
 AUTHORS: Wallace, M.R. et al.  
 TITLE: Type I Neurofibromatosis Gene: Identification  
 JOURNAL: Science  
 VOLUME: 249  
 ISSUE: 07/13/90  
 PAGES: 181-186  
 DATE: 07/13/1990  
 RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818  
 US-08-449-933-2  
 RESULT 14  
 Query Match 13.6%; score 67; DB 2, Length 2818;  
 Best Local Similarity 26.5%; Pred. No. 17; Matches 26; Conservative 18; Mismatches 32; Indels 22; Gaps 4;  
 Db 708 SCFRHICEDAEIRCAVDEVSYHNLPNNTMFEPASVSNMNSTGRALQRVMALLR-RI 766  
 Qy 62 QSIMGGSGEAMAVLVVEW-----RSLSAEGQA 88  
 Db 767 EHPTAGTNEAWEDTHAKWERQATKLILNPYPAKMEDGQA 804  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/966,049A  
 FILING DATE: 02-AUG-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Konski, Antoinette F.  
 REGISTRATION NUMBER: 34,202  
 REFERENCE/DOCKET NUMBER: 20344-20553.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 811-5600  
 TELEX: (415) 494-0792  
 TELEX: 70641 MSNFORES SFO  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2818 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 17q11.2  
 FEATURE:  
 NAME/KEY: Cleavage-site  
 LOCATION: group(583..586, 815..818, 2573..2576, 2810..2813)  
 OTHER INFORMATION: /note= "Potential CAMP-dependent"  
 OTHER INFORMATION: protein kinase recognition sites"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 2549..2556  
 OTHER INFORMATION: /note= "Potential tyrosine phosphorylation site"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: group(1264..1276, 1358..1377, 1389..1390, 1391..1392)  
 OTHER INFORMATION: /note= "Invariant residues within  
 OTHER INFORMATION: most statistically significant regions of similarity among the  
 OTHER INFORMATION: GAP family of proteins"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: group(1264..1290, 1345..1407, 1415..1430)  
 OTHER INFORMATION: /note= "Most statistically significant regions of similarity among the GAP family of  
 OTHER INFORMATION: proteins"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 496  
 OTHER INFORMATION: /note= "At variance with previously published sequence which shows an ATG methionine codon rather than an ATA isoleucine codon"  
 OTHER INFORMATION:  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 1183  
 OTHER INFORMATION: /note= "At variance with previously published sequence. Shows an CGC leucine codon rather than th  
 OTHER INFORMATION: previously published CTC"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 1555  
 OTHER INFORMATION: /note= "At variance with previously published sequence. Lacks an extra CAT histidine codon after  
 OTHER INFORMATION: this residue"  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: (2771..2772)  
 OTHER INFORMATION: /note= "Position of an 18 amino acid insertion (SQ ID NO:10) representing an alternatively  
 OTHER INFORMATION: spliced product"  
 FEATURE:

RESULT 15  
 US-08-882-704A-2  
 Sequence 2, Application US/08882704A  
 Patent No. 5879306  
 GENERAL INFORMATION:  
 APPLICANT: Jefferson, Richard A.  
 APPLICANT: Wilson, Katherine J.  
 APPLICANT: Leader, Michael  
 TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington

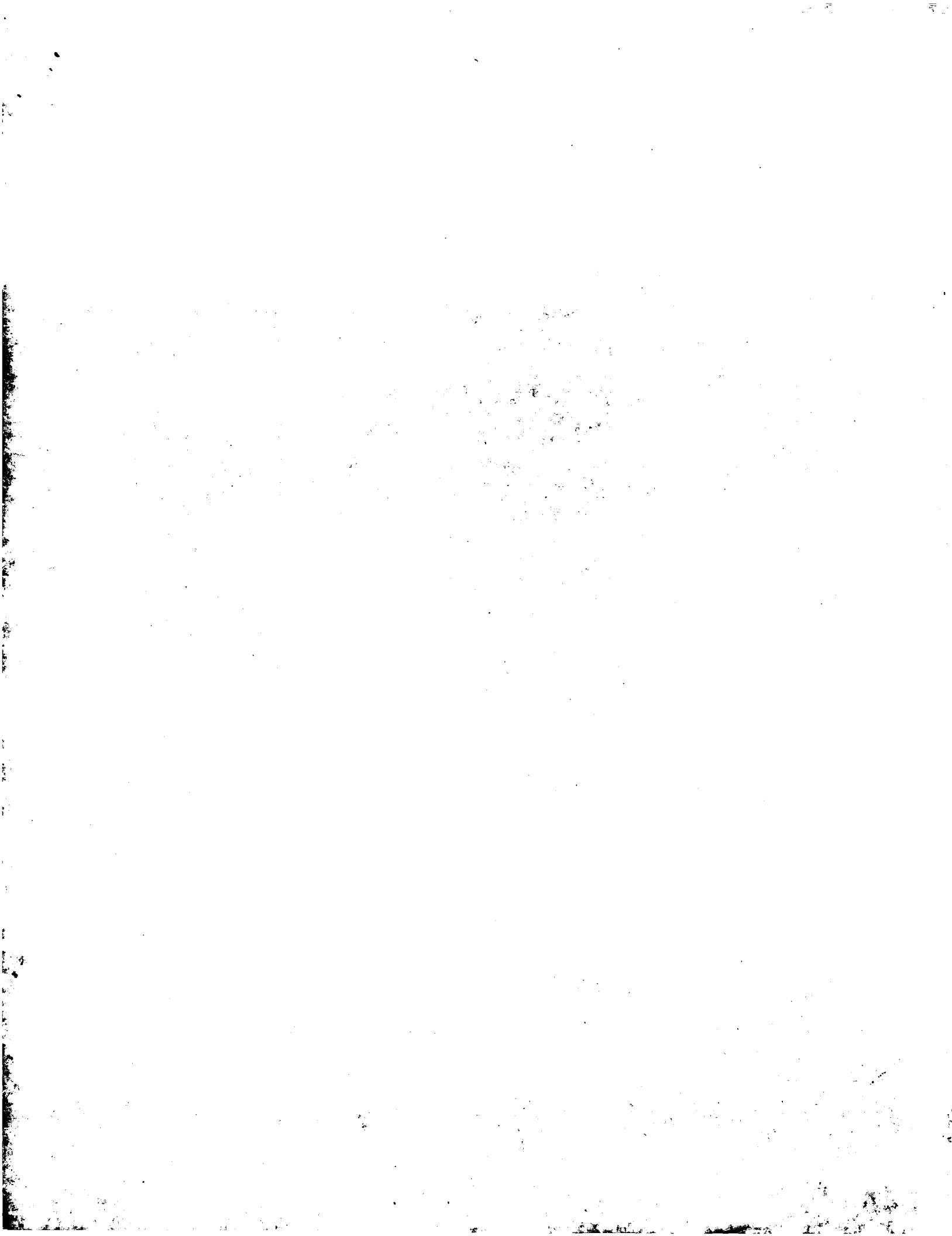
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,704A  
FILING DATE: 25-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 56799005enburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 190106.404  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 622-5031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-882-704A-2

Query Match 13.3%; Score 65.5; DB 2; Length 195;  
Best Local Similarity 26.4%; Pred. No. 0.55; Gaps 3;  
Matches 23; Conservative 17; Mismatches 34; Index 13;

QY 2 LASAVGIOQSISIFHFKSKDILRAVMEETHYNTAMMRASLEASTVRVRLIRCEL 61  
Db 36 ICKSCAISPGLYHHFPISKERALIQAILQLQDQERALARFREPSEGTHFDYM----- 87

QY 62 QSIMGGSGEAMA - VLVYERSLSRSG 85  
Db 88 ESIVSLTHEAFGQRALVVE---IMAEQ 111

Search completed: March 4, 2002, 20:24:35  
Job time: 83 sec





Query Match 22.4%; Score 110.5; DB 2; Length 190;  
 Best Local Similarity 27.0%; Pred. No. 0.00016; Gaps 3;  
 Matches 27; Conservative 25; Mismatches 35; Indels 13; C:Species: Deinococcus radiodurans  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C;Accession: F75281  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S., M.; Shen, M.; Yamatievev, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalesski, C.M.; Smith, H.O.; Venter, J.C.; Fraser, C.M.; Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: AY5250; MUID:20036896  
 A;Accession: F75281  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-197 <WHI>  
 A;Cross-references: GB:AE002068; GB:AE00513; NID:96460186; PIDN:AAF11921.1; PID:9646019  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR2376  
 A;Map position: 1

RESULT 3  
 F75281 transcription regulator, tetr family - Deinococcus radiodurans (strain R1)

Query Match 22.4%; Score 110.5; DB 2; Length 197;  
 Best Local Similarity 27.5%; Pred. No. 0.00017; Gaps 2;  
 Matches 28; Conservative 23; Mismatches 42; Indels 9; Gaps 2;  
 C:Species: Bacillus subtilis  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C;Accession: F69985  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.;Bion, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.;Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Galiech, J.; Harwood, C.R.; Hennet, A.; Hilbert, H.; Hollsappel, S.; Hosono, S.; Hulio, M; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidoz, A.; Lardino, A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y., M.; Ogawa, K.; Ogihara, A.; Ono, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Salo, T.; Scani A; Authors: Schleich, S.; Schroeter, R.; Scuffone, F.; Sekiguchi, J.; Setsuka, A.; Sei Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsit, P.; Togino, A.; Tosato, V.; Uchiya T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A; Authors: Yoshikawa, H.F.; Zunstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A;Reference number: A65580; MUID:98044033  
 A;Accession: F69985  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-194 <GBN>  
 A;Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14815.1; PID:926353  
 A;Experimental source: strain 168  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: ysia  
 C;Superfamily: Bacillus subtilis probable transcription regulator yrhi

RESULT 4  
 F75281 transcription regulator, tetr family - Deinococcus radiodurans (strain R1)

Query Match 18.8%; Score 92.5; DB 2; Length 194;  
 Best Local Similarity 27.4%; Pred. No. 0.014; Gaps 1;  
 Matches 23; Conservative 22; Mismatches 38; Indels 1; Gaps 1;  
 C:Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C;Accession: C70604  
 R:Coyle, S.T.; Brooch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Gentles, S.; Hamlin, N.; Holroyd, S.; ReJandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-54, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: AY0500; MUID:98295987  
 A;Accession: C70604  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-200 <COL>  
 A;Cross-references: GB:Z9274; GB:AL123456; NID:93261729; PIDN:CAB07159.1; PID:e306697;  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv355C  
 A;Gene: Rv355C

RESULT 5  
 F69985 transcription regulator TETR/ACRR family homolog ysia - Bacillus subtilis

Query Match 19.2%; Score 94.5; DB 2; Length 200;  
 Best Local Similarity 26.9%; Pred. No. 0.00866; Gaps 1;  
 Matches 21; Conservative 20; Mismatches 34; Indels 3; Gaps 1;  
 Matches 21; Conservative 20; Mismatches 34; Indels 3; Gaps 1;  
 C:Species: Bacillus halodurans  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
 C;Accession: F84019  
 R:Itakami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
 A;Reference number: A83650; MUID:20263314  
 A;Accession: F84019  
 A;Status: preliminary  
 A;Molecule type: DNA



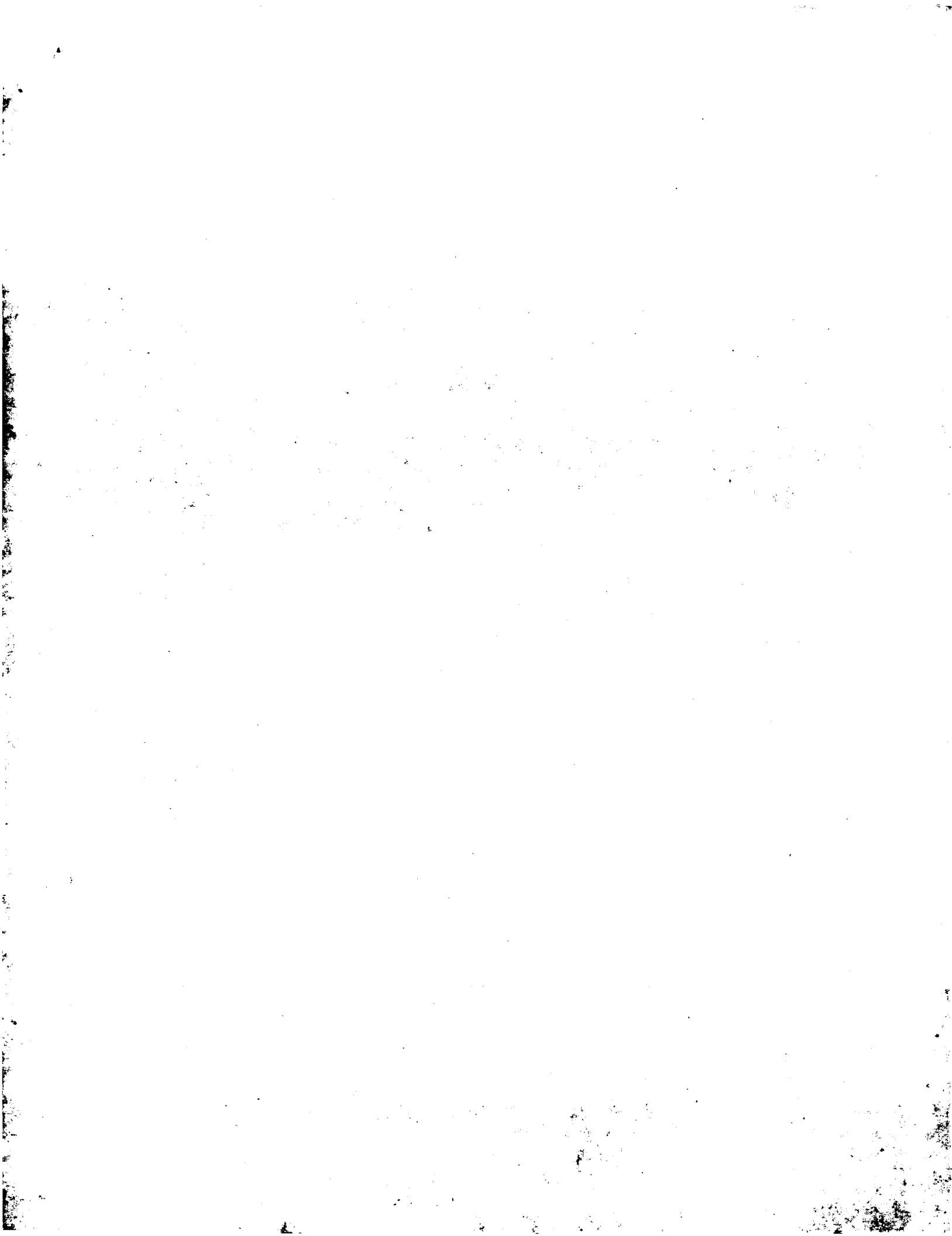


Tue Mar 5 07:57:18 2002

Db 91 AARLFQDDSIGCLMGWAVD----ASYGRSELMMA 120

H83274  
probable transcription regulator PA2957 [imported] - *Pseudomonas aeruginosa* (strain PAO1  
C; Species: *Pseudomonas aeruginosa*  
C; Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C; Accession: H83274  
A; Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A; Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A; Reference number: AB2950; PMID: 20437337  
A; Accession: H83274  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-212 <STO>  
A; Cross-references: GB:AE004722; GB:AE004091; NID:g9949054; PIDN:AAG06345.1; GSPDB:GN001  
A; Experimental source: strain PA01  
A; Genetics:  
A; Gene: PA2957

Search completed: March 4, 2002, 20:25:06  
Job time: 94 sec





		F <sup>T</sup> SEQUENCE	DNA_BIND SEQUENCE	31 195 AA;	50 21815 MW;	H-T-H MOTIF (BY SIMILARITY) B52A7/B3A605E354 CRC64;
Qy	1	DLSAVGVIQSGSIFHFRSKDEILRAVMEETHYNTAMMRASLEASTVRVLALRCE	60			
Qy	34	ELARTAGVSPALAHYFGSKDEILRAVMEETHYNTAMMRASLEASTVRVLALRCE	93			
Db	61	LQSIMGGSGEAMAVLVEWRSLSAEGQ	87			
Qy	94	FR---ADQFAPETVRAWLAFYSEAQ	115			
RESULT	2	BETI_ECOLI	BETI_ECOLI	STANDARD; PRT; 195 AA.	PRT; 195 AA.	
AC	01-AUG-1990 (Rel. 15, Created)					
DT	01-AUG-1990 (Rel. 15, Last sequence update)					
DT	20-AUG-2001 (Rel. 40, Last annotation update)					
DE	REGULATORY PROTEIN BETI.					
CN	BETI OR B0313.					
OC	Escherichia coli.					
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;					
OC	Escherichia.					
OX	NCBI_TaxID=562;					
RN	[1]	SEQUENCE FROM N.A.				
RC	STRAIN=K12;					
RX	MEDLINE=92055800; PubMed=1956285;					
RA	Lamark T., Kaasen E., Eshoo M.W., Falkenberg P., McDougall J.,					
RA	Strom A.-R.,					
RT	"DNA sequence and analysis of the bet genes encoding the					
RT	osmoregulatory choline-glycine betaine pathway of Escherichia coli. ";					
RL	Mol. Microbiol. 5:1049-1064(1991).					
RN	[2]	SEQUENCE FROM N.A.				
RP	STRAIN=K12 / MG1655;					
RX	MEDLINE=9726617; PubMed=9278503;					
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,					
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,					
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,					
RA	Mau B., Shao Y.,					
RT	The complete genome sequence of Escherichia coli K-12. ";					
RL	Science 277:1453-1474(1997).					
RN	[3]	SEQUENCE FROM N.A.				
RP	Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,					
RA	Federerle N., Hyman R., Kalman S., Kompp C., Kurdi O., Lew H.,					
RA	Liu D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;					
RL	Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.					
CC	-- FUNCTION: REPRESSOR INVOLVED IN CHOLINE REGULATION OF THE BET					
CC	- - PATHWAY: OSMOREGULATOR CHOLINE-GLYCINE BETAINE PATHWAY.					
CC	- - SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL					
CC	REPRESSORS.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">ht tp://www.isb-sib.ch/announce/</a>					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
CC	EMBL; X52905; CAA37091.1; -.					
DR	EMBL; AE00138; AAC73416.1; -.					
DR	EMBL; U73857; AAC18039.1; ALT_INIT.					
DR	PIR; S10899; S10899.					
DR	S15180.					
DR	EcoGene; EG01011; betI.					
DR	InterPro; IPR001647; HTH_Tetr.					
DR	PFAM; PF00440; tetr. 1.					
DR	PROSITE; PS00101; HTH_TETR_FAMILY; 1.					
KW	Transcription regulation; DNA-binding; Repressor; trans-acting factor;					
KW	Complete proteome.					
DR	EMBL; AE007157; AAC47851.1; -.					
DR	TIGR; MT3513; -.					
DR	Tuberculist; Rv3405C; -.					
DR	InterPro; IPR001647; HTH_Tetr.					
DR	PFAM; PF00440; tetr. 1.					
DR	PRINTS; PR00455; HTHTETR.					
KW	Hypothetical protein; Complete proteome.					



FT	CONFLICT	69	70	ML -> IV (IN REF. 2).
FT	CONFLICT	74	75	MR -> IG (IN REF. 2).
SQ	SEQUENCE	215 AA:	24404 MN:	ACA9739EA947E256 CRC64;
QY	1	DLASAVGIOSGSIFHHFSKDEILRAVMEETIHYNTAMRASLEE 45		
Db	36	EVAREAGTAPTSYRHFRDVDEGLTWDE---SGLMRLQMRRQA--RQRIAKGGSVI 88		
QY	58	RCELQSTM---GGSGEAMAVLYVWRSLSAEGAOHV 90		
Db	89	RTSVMTEFIGNNPNAFLRLRERSGTSAAFRAV 124		
RESULT	6			
Y931_HAETN				
ID Y893_HAETN				
STANDARD:			PRT:	187 AA.
AC PA4923;				
DT 01-NOV-1995 (Rel. 32, Created)				
DT 01-NOV-1995 (Rel. 32, Last sequence update)				
DT 20-AUG-2001 (Rel. 40, Last annotation update)				
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR HI0893.				
GN OS				
OC Haemophilus influenzae.				
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;				
OC Haemophilus.				
OX NCBI_TaxID=727;				
RN {1}				
RP SEQUENCE FROM N.A.				
RC STRAIN=R0 / KW20 / ATCC 51907;				
RX MEDLINE=95350630; PubMed=7542800;				
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Weirick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geohagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Ra Venter J.C.;				
RT "whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";				
RT Science 269:496-512(1995).				
RL CC				
CC REGULATORS.				
CC				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement or send an email to license@ibsb.sib.ch).				
CC				
DR EMBL; M32139; AAA0208.1; PIR; A45515; A45515.				
DR HSSP; PI9120; INGJ.				
DR InterPro; IPR001033; HSP70.				
DR Pfam; PF00012; HSP70; 1.				
DR PROSITE; PS00097; HSP70_1; 1.				
DR PROSITE; PS00329; HSP70_2; 1.				
DR PROSITE; PS01036; HSP70_3; 1.				
KW ATP-binding; Heat shock; Multigene family				
SQ SEQUENCE 676 AA; 73694 MN: AD67B783124B785E CRC64;				
Query Match	14.2 %	Score 70;	DB 1;	Length 676;
Best Local Similarity	22.9 %;	Pred No. 7.1;		
Matches	32;	Conservative	19;	Mismatches 41; Indels 48; Gaps 5;
QY 1	DLASAVGIOSGSIFHHFSK-----DEILRAVMEETI-----HYNTAMRAS 42			
Db 246	DIRNRYGIEQQLSKQMKLSKRSRCEEVKKRVLSHTRBIALDGLLPDGEYVYLKLTAR 305			
QY 43	LEEFAT_VREAFVLLIRELPS-----IMGGSGEAMAVLY----- 77			
Db 305	LEELTKIFARCLSLVQRALKDASMVKVEDIDVVLVGSSSRIPAVQALRELFRGKQLCS 365			
QY 78	-----EWRSLSAEQCAHVA 92			
Db 366	SVHPDEAVGAQAMQHVL 385			
RESULT	8			
NH34_CAEEL				
ID NH34_CAEEL				
AC Q2106;				
DT 30-MAY-2000 (Rel. 39, Created)				
DT 30-MAY-2000 (Rel. 39, Last sequence update)				
DE 30-MAY-2000 (Rel. 39, Last annotation update)				
DE NUCLEAR HORMONE RECEPTOR FAMILY MEMBER NHR-34.				
OS NHR-34 OR F58G6.5.				
Caenorhabditis elegans.				
Query Match	14.2 %;	Score 70;	DB 1;	Length 187;
Best Local Similarity	36.4 %;	Pred No. 1.6;		
Matches	16;	Conservative	11;	Mismatches 13; Indels 4; Gaps 1;



DT	03	Dynein Heavy Chain, cytosolic (DyHC).
OS	Ceutorhabditis elegans	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OX	NCBI_TaxID:6239;	
RN	[1] SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.	
RC	Strain Bristol N2;	
RX	Medline=96114101; Pubmed=8674131;	
RA	Lye R.J.; Wilson R.K.; Waterston R.H.;	
RT	"Genomic structure of a cytoplasmic dynein heavy chain gene from the nematode <i>Ceutorhabditis elegans</i> ";	
RL	Cell Motil. Cytoskeleton 32:26-36(1995).	
CC	-!- FUNCTION: dynein has ATPase activity. Cytoplasmic dynein acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules.	
CC	-!- SUBUNIT: consists of at least two heavy chains and a number of intermediate and light chains.	
CC	-!- SUBCELLULAR LOCATION: cytoplasmic.	
CC	-!- SIMILARITY: belongs to the dynein heavy chain family.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	
CC	DR	
EMBL; X61206; CAA4354.1; -.	DR	
SMART; SM00487; AAA; 1.	DR	
InterPro; IPR00393; DEAD.	DR	
InterPro; IPR00629; DEAD_ATP_helcase.	DR	
InterPro; IPR01650; Helicase_C.	DR	
Pfam; PF00270; DEAD; 1.	DR	
Pfam; PF00271; helicase_C; 1.	DR	
SMART; SM00490; HELICC; 1.	DR	
PROSITE; PS00039; DEAD_ATP_HELICASE; FALSE_NEG.	DR	
KW	Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding; DNA-binding; Helicase; Multigene family.	
FT	NP_BIND	
FT	ATP (BY SIMILARITY).	
FT	62	
FT	69	
FT	ATP	
SQ	SEQUENCE	
Query Match	13.8%; Score 68; DB 1; Length 391;	
Best Local Similarity	20.8%; Pred. No. 6.1;	
Matches	21; Conservative 27; Mismatches 39; Indels 14; Gaps 2;	
QY	7 GIQSISIFHFRSKDEILRAYMEETHYNTAMRSL-----FESTARVELV 54	
Db	13 GVEPLASFAEMGIKDDLLRSVYQGEKPSIQQRAVLPILISGRDVIAQ-QSGTGKTSMI 72	
QY	55 ALIRECLOSTINGGSGEAMAVLYVWRSLSAEGQAHVLALRD 95	
Db	73 ALTVCOI--VDTKSSEVQALILSPTRELAAQTEKVLAIGD 111	
RESULT	1.2	
DYHC_CAEEL	STANDARD; PRT; 4568 AA.	
ID	DYHC_CAEEL	
AC	Q19020;	
DT	01-NOV-1997 (Rel. 35, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
RESULT	1.3	
TGMR_STRGA	STANDARD; PRT; 226 AA.	
ID	TGMR_STRGA	
AC	P29885;	
DT	01-FEB-1995 (Rel. 31, Created)	
DT	01-FEB-1995 (Rel. 31, Last sequence update)	
DT	01-FEB-1995 (Rel. 32, Last annotation update)	
DE	TETRACENOMYCIN C TRANSCRIPTIONAL REPRESSOR.	
GN		

OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX	NCBI_TAXID=1907;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=GLA_0;
RX	MEDLINE=92276347; PubMed=1598019;
RA	Guilfoile P.G., Hutchinson C.R.;
RT	"Sequence and transcriptional analysis of the Streptomyces glaucescens tcmR tetracenomycin C resistance and repressor gene loci".
RL	J. Bacteriol. 174:3651-3658(1992).
RN	[2]
RP	FUNCTION:
RX	MEDLINE=92276348; PubMed=1592820;
RA	Guilfoile P.G., Hutchinson C.R.;
RT	"The Streptomyces glaucescens tcmR protein represses transcription of the divergently oriented tcmR and tcmA genes by binding to an intergenic operator region.";
RL	J. Bacteriol. 174:3659-3666(1992).
CC	-!- FUNCTION: REPRESSION OF THE DIVERGENTLY ORIENTED TCMR AND TCMA (TETRACENOMYCIN C RESISTANCE AND EXPORT) GENES BY BINDING TO AN INTERGENIC OPERATOR REGION. THIS BINDING IS INHIBITED BY TETRACENOMYCIN C.
CC	-!- PATHWAY: BIOSYNTHESIS OF THE POLYKETIDE ANTIBIOTIC TETRACENOMYCIN C.
CC	-!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).
CC	-----
CC	SEQUENCE FROM N.A. ( ISOFORMS A-G ).
RC	STRAIN=Heart;
RX	MEDLINE=96258425; PubMed=8777431;
RA	Kim K.K., Rajk L., Wang H., Field L.J.;
RT	"Cloning, developmental expression and evidence for alternative splicing of the murine tuberous sclerosis (TSC2) gene product.";
RL	Cell. Mol. Biol. Res. 41:515-526(1995).
RN	[2]
RP	SEQUENCE FROM N.A. AND ALTERNATIVE SPlicing.
RX	MEDLINE=96430093; PubMed=883243;
RA	Ollison P.G., Schofield J.N., Edwards Y.H., Frischauf A.M.;
RT	"Expression and differential splicing of the mouse TSC2 homolog.";
RL	Mamm. Genome 7:212-215(1995).
RN	[3]
RP	SEQUENCE OF 1-199 FROM N.A.
RC	STRAIN=BALB/C; TISSUE=Leukocyte;
RX	MEDLINE=96417643; PubMed=9743625;
RA	Sarker A.H., Imeda S., Nakano H., Terato H., Ide H., Imai K., Akiyama K., Tsutsui K., Bo Z., Kubo K., Yamamoto K., Yasui A., Yoshida M.C., Seiki S.;
RT	"Cloning and characterization of a mouse homologue (mNTBH1) of Escherichia coli endonuclease III.";
RL	J. Mol. Biol. 282:761-774(1998).
RN	[4]
RP	SEQUENCE OF 119-1805 FROM N.A.
RX	MEDLINE=20051917; PubMed=10584558;
RA	Kleymenova E.V., Declue J.E., Walker C.L.;
RT	"Genetic variants of the tuberous sclerosis 2 tumour suppressor gene in mouse t haplotypes.";
RL	Genet. Res. 74:139-144(1999).
CC	-!- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR. MAY HAVE A FUNCTION IN VESICULAR TRANSPORT, BUT MAY ALSO PLAY A ROLE IN THE REGULATION OF CELL GROWTH ARREST AND IN THE REGULATION OF TRANSCRIPTION MEDiated BY STEROID RECEPTORS. INTERACTION BETWEEN HAMARTIN AND TUBERIN MAY FACILITATE VESICULAR DOCKING. SPECIFICALLY, STIMULATES THE INTRINSIC GAPASE ACTIVITY OF THE RAS-RELATED PROTEIN RAP1 AND RAB5, SUGGESTING A POSSIBLE MECHANISM FOR ITS ROLE IN REGULATING CELLULAR GROWTH (BY SIMILARITY).
CC	-!- SUBUNIT: INTERACTS WITH HAMARTIN. MAY ALSO INTERACT WITH THE ADAPTER MOLECULE RABAPTIN 5. THE FINAL COMPLEX CONTAINS TUBERIN AND RABAPTIN 5 LINKED TO RAB5 (BY SIMILARITY).
CC	-!- ASSOCIATION WITH MEMBRANES (BY SIMILARITY).
CC	-!- ALTERNATIVE PRODUCTS: 7 ISOFORMS; A, B, C, D, E, F AND G (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPlicing.
CC	-!- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC	-!- SIMILARITY: CONTAINS 1 RAP/RAN-GAP DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).
CC	-----
ID	TSC2_MOUSE
PRT	STANDARD; PRT; 1814 AA.
AC	061037; 051007; P97723; P97724; P97725; P97727; Q9WUF6;
DT	20-AUG-2001 (Rel. 4.0, Created)
DT	20-AUG-2001 (Rel. 4.0, Last sequence update)
DT	20-AUG-2001 (Rel. 4.0, Last annotation update)
DE	TUBERTIN (TUBEROUS SCLEROSIS 2 HOMOLOG PROTEIN).
GN	TSC2.
RESULT	14
TSC2_MOUSE	STANDARD; PRT; 1814 AA.
ID	061037; 051007; P97723; P97724; P97725; P97727; Q9WUF6;
AC	061037; 051007; P97723; P97724; P97725; P97727; Q9WUF6;
DT	20-AUG-2001 (Rel. 4.0, Created)
DT	20-AUG-2001 (Rel. 4.0, Last sequence update)
DT	20-AUG-2001 (Rel. 4.0, Last annotation update)
DE	TUBERTIN (TUBEROUS SCLEROSIS 2 HOMOLOG PROTEIN).
Db	106 AFRAVLGRVRIDGELECGVGDGALHMMAMRLVERTPGLLAEYLRLRSEEMEGRALARIAARE 163
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; U37775; AAA86902.1; -.
DR	EMBL; U37775; AAA86901.1; -.
DR	MGD; MGII:102548; Tsc2.
DR	InterPro; IPR00331; Rap-GAP.
DR	Pfam; PF02145; Rap-GAP; 1.
KW	Anti-oncogene; Alternative splicing; Gtpase activation.
FT	VARSPLIC 79 115 MISSING (IN ISOFORM A).
VARSPLIC	534 572 MISSING (IN ISOFORM B).

FT	VARSPLIC	947	989	MISSING (IN ISOFORM F).
FT	VARSPLIC	947	990	MISSING (IN ISOFORM C).
FT	VARSPLIC	1245	1258	GHAPQVNTSATG-> RODALYVLSLSPAAAG (IN ISOFORM E AND ISOFORM F).
FT	VARSPLIC	1271	1293	MISSING (IN ISOFORM D, ISOFORM E AND ISOFORM F).
FT	VARSPLIC	1693	1728	GPACKCENWQKPGEVVWALPVMLTETVTLICLHQ-> M EGIVLTSVAKDNLNSVQAROMHLAN (IN ISOFORM E AND ISOFORM F).
FT	VARSPLIC	1776	1814	MISSING (IN ISOFORM F).
FT	CONFLICT	161	161	MISSING (IN REF. 2).
FT	CONFLICT	269	269	G -> S (IN REF. 2).
FT	CONFLICT	462	462	A -> R (IN REF. 2).
FT	CONFLICT	476	476	I -> N (IN REF. 2).
FT	CONFLICT	549	549	D -> N (IN REF. 2).
FT	CONFLICT	707	707	K -> N (IN REF. 4).
FT	CONFLICT	861	862	VP -> AA (IN REF. 2).
FT	CONFLICT	1127	1127	R -> P (IN REF. 2).
FT	CONFLICT	1665	1665	MISSING (IN REF. 2).
SQ	SQSEQUENCE	1814	AA;	MISSING (IN REF. 2).
SQ	SQSEQUENCE	202069	MW;	913AB7194ADA5B CRC64;

DR	Pfam: PF00440; tetrR; 1
PROSITE	PS0181; RNR_PETR_FAMILY; 1.
KW	Transcription regulation; DNA-binding; Repressor
FT	DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL)
FT	VARIANT 105 105 H->Y (IN PENICILLIN-RESISTANT
FT	ISOLATES
SQ	SEQUENCE 210 AA: 24192 MW: 41E26446CBEF57F2 CRC64:
Query Match	13 6%; Score 67; DB 1; Length 210;
best Local Similarity	16 6%; Pred No. 3-8;
Matches	24; Conservative 25; Mismatches 32; Indels 64; Gaps 4;
OY	1 DLASAVGIGOSGSIHHPKSKDELLRAVME----- - 29
Db	35 EIAQAGAVGVRGALYWHFRKRNKEKDFDALFORCIDDIENCTAQDADAEGGSWTVRHTLH 94
Qy	30 -----EITY----- -----NPAMMASLEASTVRVLALLI--RCLEQSIM 65
Db	95 FFERLQSNDIHYKHFNHLFLKCETEQNAMMIAVARKQIAWIKETIVTAEVENDLA 154

Query Match 13.7%; Score 67.5; DB 1; Length 1814;  
 Best Local Similarity 27.5%; Pred. No. 41;  
 Matches 25; Conservative 17; Mismatches 36; Index 13; Gaps 0  
 Qy VGIQSSGIFHHRSKDEILRRAVMEETI-----HYNTAMRASLEASTVRERVLALLR 58  
 Db 513 VDLAEGGHTHHFNLSLDDIEKWMARSLSPPPELEERDVLAVHSASLEDVKTAVLGLVLQL 572  
 Qy CBLQSINGGSGBMAVLYEWRLSAEBOGAH 89  
 Db 573 TRLYTLPA---SHATRYEE-SLISHIQHL 597

Search completed: March 4, 2002, 20:28:54  
Job time: 97 sec

RESULT 19

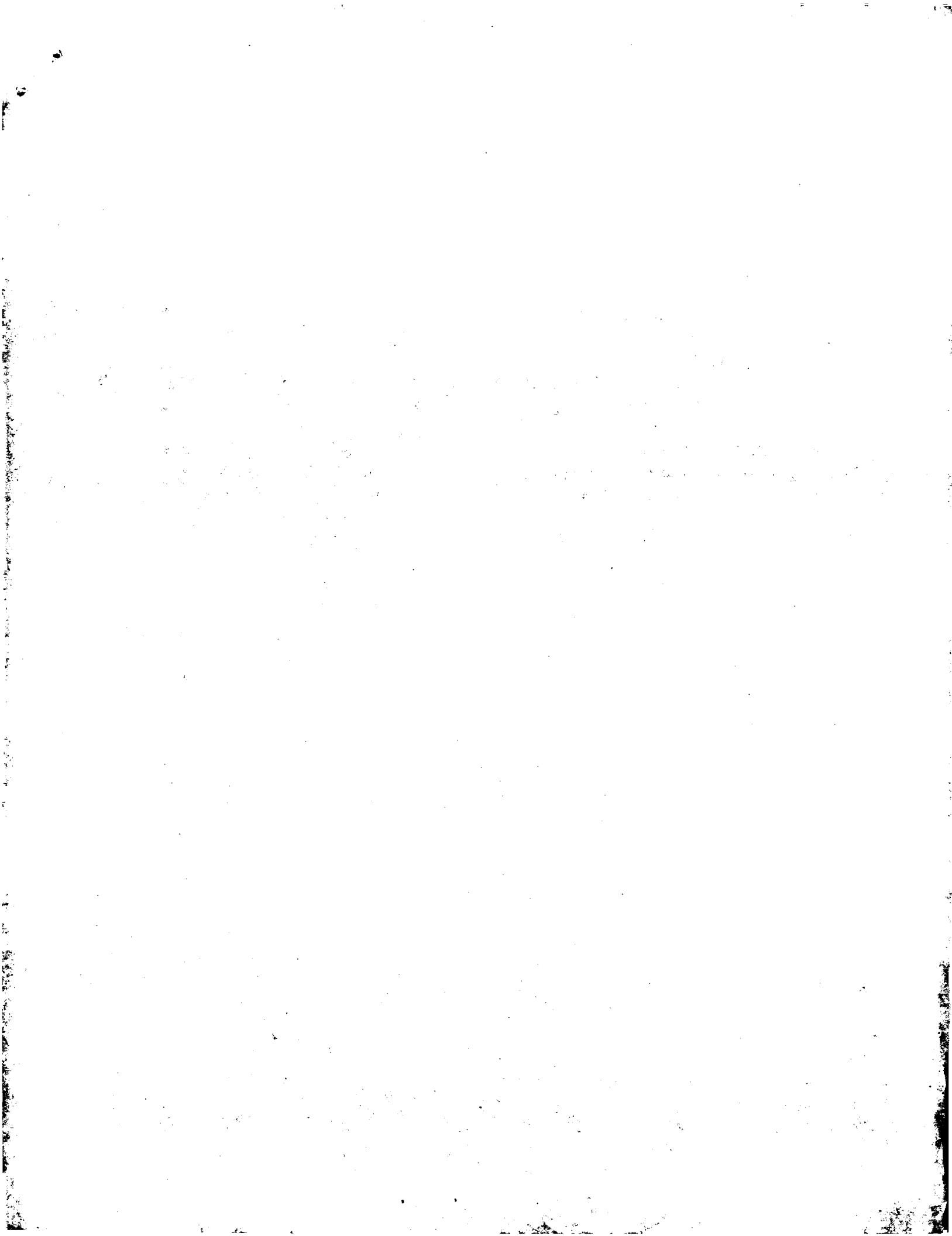
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
[1]  
SEQUENCE FROM N. A.  
STRAIN=PA19;  
MEDLINE=91254732; PUBMED=8196548;  
Pan W., Spratt B.G.;  
"Regulation of the permeability of the gonococcal cell envelope by  
the mtr system";  
Mol. Microbiol. 11:769-775(1994).  
-1- FUNCTION: PUTATIVE REPRESSOR OF MTRC GENE. CONTROLS THE  
PERMEABILITY OF THE CELL ENVELOPE TO HYDROPHOBIC COMPOUNDS SUCH AS  
ANTIBIOTICS AND DETERGENTS.  
-1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
REGULATORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

DR EMBL; 225797; CAA81047.1; -.  
 DR PIR; S40250; S40250.  
 DR HSSP; P09164; 2TRT.  
 DR InterPro; IPR001647; HTH\_TETR.

Tue Mar 5 07:57:19 2002

us-09-966-608-2.rsp



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 4, 2002, 20:25:12 ; Search time 27.32 Seconds

(without alignments)  
535.404 Million cell updates/sec

Title: US-09-966-608-2

Perfect score: 493

Sequence: 1 DLASAVGIGSGSIFHHFKSK..... SLSAQAHVLAIRDVYEQI 100

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: SPTRMEL\_17:\*
- 2: sp-archaea:\*
- 3: sp-bacteria:\*
- 4: sp-fungi:\*
- 5: sp-human:\*
- 6: sp-invertebrate:\*
- 7: sp-mammal:\*
- 8: sp-organelle:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

RESULT	1	PRELIMINARY:	PRT:	207 AA.
ID	Q9RFR4			
AC	Q9RFR4;			
DT	01-MAY-2000	( TREMBLrel. 13, Created)		
DT	01-JUN-2001	( TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2001	( TREMBLrel. 17, Last annotation update)		
DE	PUTATIVE TRANSCRIPTIONAL REGULATOR PR12.			
OS	Pseudomonas fluorescens.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas			
OX	NCBI_TaxID=294;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20349410; PubMed=10889151;			
RA	Sutton C.L., Kim J., Yamane A., Dallwadi H., Wei B., Landers C., Tagan R., Braun J.; "Identification of a novel bacterial sequence associated with Crohn's disease.",			
RT				
RL	Gastroenterology 119:23-31(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Wei B., Huang T., Dallwadi H., Sutton C.L., Braun J.; "12, a Crohn's disease-associated microbial gene and T-cell superantigen, is a species-specific product of Pseudomonas fluorescens.", Submitted (MAR-2001) to the EMBL/genBank/DBJ databases.			
CC	-1- SIMILARITY TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.			
DR	EMBL: AF173683; AAK16600; 1; -.			
DR	Inter-Pro: IPRO01647; HTH_Tetr.			
DR	PFam: PF00440; tetr; 1.			
DR	PRINS; PR005; HHTTETR.			
DR	PROSTE: PS01081; HTH_TETR_FAMILY: 1.			
KW	RNA binding; transcription regulation; 2			
SEQUENCE	207 AA; 231.07 MW; DDJAFB7FE10A736 CRC64;			

Query Match	98.2%	Score	484	DB	2	Length	207
Best Local Similarity	98.0%	Pred. No.	2.7e-42				
Matches	98	Conservative	1	Mismatches	1	Indels	0
						Gaps	0
QY	1	DLASAVGIGSGSIFHHFKSKDELRVMEETIHYNTAMMRASLEASTVRERVLALIRCE	60				

1	484	98.2	207	2	Q9RFR4	20	80	16.2	186	2	Q9I497
2	429	87.0	198	2	Q9H2W2	21	80	16.2	212	2	Q9HZP1
3	110.5	22.4	190	2	Q95306	22	79	16.0	209	2	Q9XA31
4	110.5	22.4	197	2	Q9RRV9	23	78	15.8	196	2	Q9RMF4
5	94.5	19.2	200	2	P96839	24	78	15.8	211	2	Q9A905
6	92.5	18.8	194	2	P94548	25	78	15.8	217	2	Q9K3M6
7	92.5	18.8	215	2	Q9K8P5	26	77.5	15.7	187	2	Q9K017
8	90	18.3	192	2	Q9AK88	27	77.5	15.7	194	2	Q9S261
9	90	18.3	222	2	Q9SL34	33	76.5	15.5	202	2	Q9A250
10	89	18.1	192	2	Q96727	34	76.5	15.5	213	2	Q9JV29
11	86	17.4	192	2	Q9K3Q4	35	76.5	15.5	217	2	Q87854
12	86	17.4	195	2	Q9K8A4	36	76	15.4	192	2	Q9ZGB7
13	85	17.2	197	2	Q9HTJ0	37	76	15.4	195	2	Q9X8M4
14	84.5	17.1	215	2	Q9S253	38	76	15.4	202	2	Q9KXT8
15	84	17.0	206	2	Q9X7X0	39	76	15.4	213	2	Q9S250
16	83	16.8	214	2	Q9JNB9	40	75.5	15.4	215	2	Q9AA74
17	82.5	16.7	223	2	Q9X9X9	41	75.5	15.3	209	2	Q9S2V6
18	82	16.6	265	2	Q9AA36	42	75	15.2	190	2	Q07388
19	81	16.4	190	2	Q32228	43	75	15.2	193	2	Q9A199
						44	75	15.2	192	2	Q912Q9
						45	75	15.2	1489	5	Q9WIR3

## ALIGNMENTS

			CC	-1- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
Db	49	DLASAVGIOGSFFHRSKDEILRLRAVMEETHYNTAMRASLEASTVRERVLALIRCK	DR	EMBL; 22835; CAN8221L.1; -.
Qy	61	LOSIMGGSEAMAWLVWEWSLSAEGQAHVALRDVEQI	DR	InterPro; IPR01647; HTH_TetR.
Db	109	LOSIMGGSEAMAWLVWEWSLSAEGQAHVALRDVEQI	DR	Pfam; PF00440; tetr; 1.
GN	P2885.	Pseudomonas aeruginosa.	DR	PROSITE; PS01081; HTH_TETR_FAMILY; 1.
OS	OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;	KW	DNA-binding; Ribosomal protein; Transcription regulation.
OC	OC	Pseudomonas aeruginosa.	SQ	SEQUENCE 190 AA; 21692 MW; 0C6ED92A4BAA69C0 CRC64;
OX	OC	Pseudomonas aeruginosa.		
RA	NCBI_TAXID=287;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PAOL;			
RX	MEDLINE=2043/337; PubMed=1094043;			
RA	Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalki D.J., Laiou M., Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Pulsifer I.T., Reitzer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.; RT "complete genome sequence of <i>Pseudomonas aeruginosa</i> PAOL, an opportunistic pathogen.";			
RT	Nature 406:959-960(2000).			
RL				
DR	EMBL; AE004714; AAC06273.1; -.			
DR	InterPro; IPR001647; HTH_Tetr.			
DR	PRINTS; PR00455; HTHTEPR.			
DR	PROSITE; PS01081; HTH_TETR_FAMILY; 1.			
KW	Complete proteome; DNA-binding; Transcription regulation.			
SQ	SEQUENCE 198 AA; 22057 MW; 079217CC7B8FF85 CRC64;			
RESULT	3			
Q9306	PRELIMINARY;			
ID	09306	PRT; 190 AA.		
AC	Q9306;			
DT	01-NOV-1996 (Tremblel. 01, created)			
DT	01-NOV-1996 (Tremblel. 01, last sequence update)			
DT	01-JUN-2001 (Tremblel. 17, last annotation update)			
DE	305 RIBOSOMAL PROTEIN S21.			
OS	Clostridium pasteurianum.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.			
OC	NCBI_TaxID=1501;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RA	Meyer J;			
RL	Submitted (Nov-1993) to the EMBL/GenBank/DBJ databases.			
RESULT	4			
Q9RV9	PRELIMINARY;			
ID	Q9RV9	PRT; 197 AA.		
AC	Q9RV9;			
DT	01-MAY-2000 (Tremblel. 13, created)			
DT	01-MAY-2000 (Tremblel. 13, last sequence update)			
DT	01-JUN-2001 (Tremblel. 17, last annotation update)			
DE	TRANSCRIPTIONAL REGULATOR, TETR FAMILY.			
DR	2376.			
GN	Deinococcus radiodurans.			
OC	Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.			
OX	NCBI_TAXID=1299;			
RA	SEQUENCE FROM N.A.			
RC	STRAIN=R1;			
RX	MEDLINE=20036896; PubMed=10567266;			
RA	White O., Elsen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson D.L., Moffat K.S., Oin H., Jiang L., Pamphile W., Crosby M., Shen M., Venkatesan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C., Fraser C.M., "Genome sequence of the radioresistant bacterium <i>Deinococcus radiodurans</i> R1";			
RT	Science 286:1571-1577/1999.			
RL				
CC	-1- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.			
DR	EMBL; AB00268; AAF11921.1; -.			
DR	TIGR; DR2375; -.			
DR	InterPro; IPR001647; HTH_Tetr.			
DR	PRINTS; PR00455; HTHTEPR.			
KW	Complete proteome; DNA-binding; Transcription regulation.			
SQ	SEQUENCE 197 AA; 22307 MW; 769F2960C5F848D CRC64;			
RESULT	5			
Q9306	PRELIMINARY;			
ID	09306	PRT; 190 AA.		
AC	Q9306;			
DT	01-NOV-1996 (Tremblel. 01, created)			
DT	01-NOV-1996 (Tremblel. 01, last sequence update)			
DT	01-JUN-2001 (Tremblel. 17, last annotation update)			
DE	305 RIBOSOMAL PROTEIN S21.			
OS	Clostridium pasteurianum.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.			
OC	NCBI_TaxID=1501;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RA	Meyer J;			
RL	Submitted (Nov-1993) to the EMBL/GenBank/DBJ databases.			
RESULT	5			
Q9639	PRELIMINARY;			
ID	Q9639	PRT; 190 AA.		
AC	Q9639;			
DT	01-NOV-1996 (Tremblel. 01, created)			
DT	01-NOV-1996 (Tremblel. 01, last sequence update)			
DT	01-JUN-2001 (Tremblel. 17, last annotation update)			
DE	305 RIBOSOMAL PROTEIN S21.			
OS	Clostridium pasteurianum.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.			
OC	NCBI_TaxID=1501;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RA	Meyer J;			
RL	Submitted (Nov-1993) to the EMBL/GenBank/DBJ databases.			

ID P96839 PRELIMINARY; PRT; 200 AA.  
 AC P96839;  
 DT 01-MAY-1997 (TREMBREL 03, last sequence update)  
 DT 01-JUN-2001 (TREMBREL 17, last annotation update)  
 DE HYPOTHETICAL 22.9 kDa PROTEIN.  
 GN RV3557C OR MTC06G11.04C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TAXID=1773;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=R37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Connar R.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Shetton S., Squares S., Squares R., Sulston J.E.,  
 RA Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence";  
 RL Nature 335:537-541(1998).  
 CC !- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
 DR EMBL: Z92774; CAB0159.1; -.  
 DR Tuberculist: RV3557C; -.  
 DR InterPro: IPR001647; HTH\_Tetr.  
 DR Pfam: PF00440; tetrR; 1.  
 KW Complete proteome; DNA-binding; Hypothetical protein;  
 KW Transcription regulation; MW: 22905 CRC64;  
 SQ SEQUENCE 200 AA; E1B29D35CF4D955F CRC64;

RESULT 6  
 P94548 PRELIMINARY; PRT; 194 AA.  
 AC P94548;  
 DT 01-MAY-1997 (TREMBREL 03, Created)  
 DT 01-JUN-1998 (TREMBREL 05, last sequence update)  
 DT 01-JUN-2001 (TREMBREL 17, last annotation update)  
 DE HYPOTHETICAL 22.0 kDa PROTEIN.  
 GN YSIA.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/staphylococcus group; Bacillus.  
 OC NCBITaxID=1423;  
 RN [1] SEQUENCE FROM N.A.  
 RP MEDLINE=97217425; PubMed=9063446;  
 RA Goethel S.F., Schmid R., Wipat A., Carter N.M., Emmerson P.T.,  
 RA Harwood C.R., Marahiel M.A.;  
 RT "An internal F506-binding domain is the catalytic core of the prolyl  
 isomerase activity associated with the *Bacillus subtilis* trigger  
 factor";  
 RL Eur J Biochem 244:59-65(1997).  
 RN [2] SEQUENCE FROM N.A.

RESULT 6  
 P94548 PRELIMINARY; PRT; 194 AA.  
 AC P94548;  
 DT 01-MAY-1997 (TREMBREL 03, Created)  
 DT 01-JUN-2001 (TREMBREL 17, last annotation update)  
 DE HYPOTHETICAL 22.0 kDa PROTEIN.  
 GN YSIA.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/staphylococcus group; Bacillus.  
 OC NCBITaxID=1423;  
 RN [1] SEQUENCE FROM N.A.  
 RP MEDLINE=97217425; PubMed=9063446;  
 RA Goethel S.F., Schmid R., Wipat A., Carter N.M., Emmerson P.T.,  
 RA Harwood C.R., Marahiel M.A.;  
 RT "An internal F506-binding domain is the catalytic core of the prolyl  
 isomerase activity associated with the *Bacillus subtilis* trigger  
 factor";  
 RL Eur J Biochem 244:59-65(1997).  
 RN [2] SEQUENCE FROM N.A.

RESULT 7  
 O9K8P5 PRELIMINARY; PRT; 215 AA.  
 ID O9K8P5  
 AC O9K8P5;  
 DT 01-OCT-2000 (TREMBREL 15, Created)  
 DT 01-OCT-2000 (TREMBREL 15, last sequence update)  
 DT 01-JUN-2001 (TREMBREL 17, last annotation update)  
 DE TRANSCRIPTIONAL REGULATOR (TETR/ACRR FAMILY).  
 GN B2958.  
 OS Bacteria; Firmicutes; Bacillus/Clostridium group;

OC BACILLUS/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=205/2582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuwahara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.";  
 RL Nucleic Acids Res. 28;:317-4331(2000).  
 CC -!- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
 DR EMBL: AP001517; BAB06677.1; -  
 DR Interpro: IPR001647; HTH\_Tetr.  
 DR Pfam: PF00440; teer; 1.  
 DR PRINTS: PR00455; HTHTER.  
 KW Complete proteome; DNA-binding; transcription regulation.  
 SQ SEQUENCE 215 AA; 25323 MW; EB9E1FFA842F3F87 CRC64;

RESULT 8  
 Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR, TETR/ACRR FAMILY.  
 GN CC1345.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Elsner J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Portnoy I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Derby R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Heft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khoruri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Yamashita J., Emrulaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005810; AAK2326.1; -  
 DR TIGR; CC1345; -.  
 KW Complete proteome.  
 SQ SEQUENCE 192 AA; 20527 MW; 041950A960691667 CRC64;

Query Match 18.3%; Score 90; DB 2; Length 192;  
 Best Local Similarity 25.9%; Pred. No. 0.11; Gaps 4;  
 Matches 28; Conservative 22; Mismatches 34; Indels 24; Gaps 4;

Q9A8K8 PRELIMINARY; PRT; 192 AA.  
 ID Q9A8K8;  
 AC 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE TRANSCRIPTIONAL REGULATOR (TETR/ACRR FAMILY).  
 GN CC1345.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquiflex.  
 OX NCBI\_TaxID=63363;  
 RN SEQUENCE FROM N.A.  
 RX STRAIN=VF5;  
 RA MEDLINE=9819666; PubMed=9537320;  
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Snead M.A., Kellie M., Aufay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RA "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus."  
 RT Nature 392:353-358(1998).  
 RL Nature 392:353-358(1998).  
 CC -!- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
 DR EMBL: AE000776; AAC0781.1; -  
 DR Interpro: IPR002106; AA\_tRNA\_ligase\_II.



SQ	SEQUENCE	197 AA:	21861 MW:	5BF9AB16F73B96C0 CRC64:	
Query Match	Best Local Similarity	17.2%	Score	85;	DB 2; Length 197;
Matches	31;	Conservative	29.8%;	Pred.	No. 0.38;
Qy	2	LASAVGIOGSFIFHFKSKDEILRRAVMEETHYNTAMMRASLEASTVRVLALI---	57	Mismatches	31;
Db	35	TARLAGVSNGTISHYFODKNGLLEATR--HILLASLSKA-----	57	Indels	26;
Qy	58	RBLQSMGG-----SGEAMAVLVEWRSLSAEGOAHVIALR	94	Gaps	5;
Db	85	RAHLRATIVEGVNFDDSQVNGPANKTWLAFWAT----SMHQPLR	123		
RESULT	14				
Q9S253		PRELIMINARY;	PRM;	215 AA.	
AC	Q9S253;				
DT	01-MAY-2000	(TREMBLREL. 13, Created)			
DT	01-MAY-2000	(TREMBLREL. 13, Last sequence update)			
DT	01-JUN-2001	(TREMBLREL. 17, Last annotation update)			
DE		PUTATIVE TRANSCRIPTIONAL REGULATOR.			
GN		SC130A_20C.			
OS		Streptomyces coelicolor.			
OC		Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC		Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
OX		[1]			
RN		SEQUENCE FROM N.A.			
RP		STRAIN=A3(2);			
RC		SEQUENCE FROM N.A.			
RC		STRAIN=A3(2);			
RA		Olive K., Harris D.;			
RA		"A set of ordered cosmids and a detailed genetic and physical map for			
RA		the 8 Mb Streptomyces coelicolor A3(2) chromosome.";			
RA		Submitted (MAR-1999) to the EMBL/GenBank/DDJB databases.			
RA		Submitted (JUL-1999) to the EMBL/GenBank/DDJB databases.			
RA		[2]			
RA		SEQUENCE FROM N.A.			
RA		STRAIN=A3(2);			
RA		SEQUENCE FROM N.A.			
RA		STRAIN=A3(2);			
RA		REDENBACH M., KIESER H.M., DENAPAITA D., EICHNER A., CULLUM J.,			
RA		KINASHI H., HOPWOOD D.A.;			
RA		"A set of ordered cosmids and a detailed genetic and physical map for			
RA		the 8 Mb Streptomyces coelicolor A3(2) chromosome.";			
RA		MOI. MICROBIOL. 21:77-96(1996).			
RA		-!- SIMILARITY: TO THE TETRACYCERR FAMILY OF TRANSCRIPTIONAL REGULATORS.			
RA		EMBL; AL049485; CAB39708.1; .			
DR		InterPro; IPR001647; RTH_TETR.			
DR		PRINTS; PR00440; tetr; 1.			
DR		PRINTS; PR00455; HTHTTR.			
KW		DNA-binding; transcription regulation.			
SQ		SEQUENCE 206 AA; 75D88R30D7476A75 CRC64;			
Query Match	Best Local Similarity	17.0%	Score	84;	DB 2; Length 206;
Matches	21;	Conservative	21.0%;	Pred.	No. 0.51;
Qy	1	DLASAVGIOGSFIFHFKSKDEILRRAVMEETHYNTAMMRASLEASTVRVLALIRCE	60	Mismatches	46;
Db	35	DYAAALGVSNALVLYHSTKERVAAFTHAEDDLAQRLGQSLARRSALRRLRAAVRW	94	Indels	8;
Qy	61	LOSIMGSGEAMAVLVEWRSLSAEGOAHVIALRDWIEO	100	Gaps	2;
Db	95	AFT---GPAKGWLWIEGWAAAPREP---ALREVTREL	126		
RESULT	15				
Q9X7X0		PRELIMINARY;	PRT;	206 AA.	
ID					

Query Match 17.1%; Score 84.5; DB 2; Length 215;  
 Best Local Similarity 23.5%; Pred. No. 0.47;  
 Matches 23; Conservative 22; Mismatches 44; Indels 9; Gaps 2;

Qy 1 DLASAVGIOGSFIFHFKSKDEILRRAVMEETHYNTAMMRASLEASTVRVLALI----- 56  
 Db 46 DIAGRAGMSPAALYIHKTKEELLHRISRIGHTRAVALRSQAEGSSAERLADAVSSP 105

Qy 57 TRCELQSIMGSGEAMAVLVEWRSLSAEGOAHVIALR 94  
 Db 106 VRWH-----AGRRTTARVVOYELDALGPAREEELTLR 138

Search completed: March 4, 2002, 20:28:29  
 Job time: 197 sec

Tue Mar 5 07:57:19 2002

us-09-966-608-2.rspt

